8-2015

Identification of Genes Required to Synthesize an Antibiotic-like Compound from the Soil Bacterium Rhodococcus sp. MTM3W5.2

Amber L. Ward
East Tennessee State University

Identification of Genes Required to Synthesize an Antibiotic-like Compound from the Soil Bacterium *Rhodococcus* sp. MTM3W5.2

A thesis

presented to

the faculty of the Department of Health Sciences

East Tennessee State University

In partial fulfillment

of the requirements for the degree

Master of Science in Biology

---

by

Amber L. Ward

August 2015

Dr. Bert C. Lampson, Chair

Dr. Eric L. Mustain

Dr. Christopher L. Pritchett

Dr. Abbas Shilabin

Keywords: *Rhodococcus*, polyketide synthase, antibiotic, natural product

ABSTRACT

Identification of Genes Required to Synthesize an Antibiotic-like Compound from the Soil Bacterium *Rhodococcus* sp. MTM3W5.2

by

Amber L. Ward

Rhodococcus is a soil bacterium, member of the *Actinobacteria*, and a close relative of the prolific small molecule producer *Streptomyces*. Recent interest in *Rhodococcus* as an under investigated source of possible bioactive secondary metabolites is sparked by the discovery of many polyketide synthase and non-ribosomal peptide synthetase genes of unknown function from sequenced *Rhodococcus* genomes. *Rhodococcus* species strain MTM3W5.2 was recently shown to produce a strong inhibitory compound with activity against most strains of *Rhodococcus* and closely related genera. A goal of this investigation is to discover the gene(s) required to synthesize this inhibitory molecule. The engineered *Rhodococcus* transposon, pTNR, was used to generate random insertional mutations in the genome of MTM3W5.2. The transposon insertion sites for 8 non-producing mutants were cloned and sequenced. Genes that encode polyketide synthases usually form parts of large biosynthetic gene clusters responsible for the production of small polyketide molecules.

DEDICATION

To mom and dad

ACKNOLWEDGEMENTS

I would like to sincerely thank Dr. Bert Lampson, my advisor and committee chair, for being an excellent mentor. Thank you for your patience, infinite support and guidance and allowing me to both succeed and fail. Thank you for allowing me to work on an amazing project and giving me the opportunity to continually bring ideas to you to further the project. To my committee members, Dr. Christopher Pritchett, Dr. Eric Mustain and Dr. Abbas Shilabin, thank you for your exceptional guidance and going beyond your duties as a committee member. Thank you Dr. Shilabin for allowing us to create a collaborative project with you and your lab. Thank you Dr. Sean Fox, Robin Grindstaff, Sean Stacey and Valeria Barisic for the technical support, letting me borrow lab equipment and lending an ear when I was frustrated and couldn't see the light at the end of the tunnel. To Danielle Williams and Spencer Cate, thank you for washing an endless number of beakers for me and happily completing other lab duties to relieve some of my stress. I would like to thank Rhesa Dykes at the ETSU Molecular Biology Core for your kindness and patience with me while I was learning how to sequence and use your facilities.

To my parents, thank you for your endless love and support. Thank you for always encouraging me to push through when I thought that I couldn't go any further and teaching me that there is never a dream that is too big. To my sister, Tiffany, and brother, Michael, thank you for being great siblings who always support me and being two of my best friends. Thank you, God, for your infinite blessings and for guiding me to a career that I can be passionate about.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Antibiotic Resistance

Antibiotic resistant infections have reached a crisis level within the last decade. There has been an emergence of new bacterial pathogens, development of antibiotic resistance, and resistance to several different classes of drug compounds simultaneously. Thus, high rates of morbidity and mortality continue to be a problem for antibiotic resistant infections. Antibiotic resistance is a bacteria's ability to resist the effects of an antibiotic through various processes such as efflux pumps, enzymatic target modification, overproducing the target, and drug resistance genes (r genes) (CDC 2013). In 2013, the CDC reported that over 2 million illnesses can be attributed to antibiotic resistant infections each year within the United States, and of these 2 million illnesses, approximately 23,000 result in death (CDC 2013). With a decline in antibiotic discovery, the growing number of new pathogens as well as known pathogens finding ways to become more resistant, microbes are winning the war on infectious diseases. Such a drastic decline in drug discovery has occurred that between 1962 and 2000, no novel antibiotic scaffolds were discovered during this period (Seyedsayamdost 2014). Since 2000, just four new classes of antibiotics have been marketed (Fischbach and Walsh 2009). These new classes have yet to make a large impact clinically, and because of this, physicians still rely on the antibiotics discovered over half a century ago, many of which are not as effective as they once were in combating pathogens. Finding new classes of antibiotics seems to be the answer to many of the problems that we face, but research shows that within two years of a

novel class of antibiotic being introduced, resistance is usually observed (Coates et al. 2011).

For many decades, antibiotic discovery was on the forefront of research. It was backed by the government, public and private institutions, and pharmaceutical companies, and in the 1960s, the United States Surgeon General, William H. Stewart, stated, "[it] is now time to close the book on infectious disease and declare the war on pestilence won" (Spellberg et al. 2008). Although an encouraging declaration, the next three decades proved such a statement to be false. With resolutions and funding being thrown at the war on pathogens and most recently, the Obama administration continuing this trend by pledging a $1.2 billion, five-year plan to combat the antibiotic resistance crisis, all has been done to no avail, so far. While the research has had its ebbs and flows, a recent database shows that bacteria continue to evolve. The database of available bacterial genomes provides information on more than 20,000 r genes of more than 400 different types (Liu and Pop 2009), and these numbers continue to grow.

Three groups of pathogens have emerged as major threats to society because of their antibiotic resistance. One is methicillin-resistant *Staphylococcus aureus* (MRSA). This group is estimated to cause approximately 19,000 deaths in the United States each year, and can lead to an extra $3 to 4 billion of additional healthcare costs aside from morbidity (Fischbach and Walsh 2009). The second class is multidrug-resistant (MDR) and pan-drug-resistant (PDR) gram-negative bacteria. These strains of bacteria such as *Escherichia coli* and *Pseudomonas aeruginosa* are less common than MRSA, but are essentially untreatable (Fischbach and Walsh 2009). The classes of drugs that are used to treat gram-negative infections do not inhibit these pathogens. In the third class, composed of MDR and

extensively-drug-resistant (XDR) strains of *Mycobacterium tuberculosis* (MDR-TB and XDR-TB). This class of pathogens has become an increasing problem, including in developed countries such as England and the United States. MDR-TB is treated with a two-year course of antibiotics that has serious side effects, and XDR-TB is even more difficult to treat and commonly fatal (Fischbach and Walsh 2009).

Recent research indicates the importance of gene exchange as a universal property of bacteria. Gene exchange has been occurring throughout microbial evolution. Of particular importance to the emergence of resistant pathogens has been the plasmid-mediated transfer of r genes (Davies and Davies 2010). Plasmid-mediated r genes commonly provide multiple ways to resist a particular antibiotic (Table 1). For example, β-lactamases are a class of enzymes that are responsible for inactivating a class of antibiotics called β-lactams. This class of antibiotics is commonly given for *Escherichia coli*, *Salmonella enterica*, and *Klebsiella pneumoniae* infections, but in the last century, there has been a strong correlation between antibiotic use and antibiotic resistance in these infections. To date, up to 1,000 β-lactamase r genes have been identified (Davies and Davies 2010).

Table 1: Modes of action and resistance mechanisms of commonly used antibiotics (Davies and Davies 2010)

Antibiotic class	Example(s)	Target	Mode(s) of resistance
β-Lactams	Penicillins (ampicillin), cephalosporins (cephamycin), penems (meropenem), monobactams (aztreonam)	Peptidoglycan biosynthesis	Hydrolysis, efflux, altered target
Aminoglycosides	Gentamicin, streptomycin, spectinomycin	Translation	Phosphorylation, acetylation, nucleotidylation, efflux, altered target
Glycopeptides	Vancomycin, teicoplanin	Peptidoglycan biosynthesis	Reprogramming peptidoglycan biosynthesis
Tetracyclines	Minocycline, tigecycline	Translation	Monooxygenation, efflux, altered target
Macrolides	Erythromycin, azithromicin	Translation	Hydrolysis, glycosylation, phosphorylation, efflux, altered target
Lincosamides	Clindamycin	Translation	Nucleotidylation, efflux, altered target
Streptogramins	Synercid	Translation	C-O lyase (type B streptogramins), acetylation (type A streptogramins), efflux, altered target
Oxazolidinones	Linezolid	Translation	Efflux, altered target
Phenicols	Chloramphenicol	Translation	Acetylation, efflux, altered target
Quinolones	Ciprofloxacin	DNA replication	Acetylation, efflux, altered target
Pyrimidines	Trimethoprim	C_1 metabolism	Efflux, altered target
Sulfonamides	Sulfamethoxazole	C_1 metabolism	Efflux, altered target
Rifamycins	Rifampin	Transcription	ADP-ribosylation, efflux, altered target
Lipopeptides	Daptomycin	Cell membrane	Altered target
Cationic peptides	Colistin	Cell membrane	Altered target, efflux

Natural Products

A natural product is a small organic chemical compound or molecule that is produced by a living organism that is found in nature. The products can be extracted from tissues, plants, microorganisms, and animals. Natural products include secondary metabolites that are commonly organism specific. These products are not essential for growth, development, or reproduction of the organism, but can be beneficial for the organism to compete in its environment. Molecules produced from the fundamental processes of the Krebs cycle, photosynthesis, and glycolysis can yield intermediates that lead to the production of secondary metabolites (Dias et al. 2012). Microorganisms live in a world where they are constantly competing with other organisms (the same or different species) for the resources available in the environment. Bacteria have evolved so that they can be successful in these environments by having the ability to produce natural products, such as siderophores, pigments, signaling molecules, and toxins. Siderophores are small high-affinity iron chelating compounds that can be released from the bacterial cell to scavenge iron from the environment and be brought back into the cell so that the iron can be used for life processes (Neilands 1995). In addition to siderophores, bacteria can produce pigments. Xanthomonadin is a pigment produced by *Xanthomonas oryzae,* and it protects the bacteria from damage due to light (Rajagopal et al. 1997). Finally, Cyanobacteria can produce secondary metabolites that are toxic and make freshwater dangerous for drinking and recreation (Penn et al. 2014). Although these secondary metabolites can be beneficial to bacteria, natural products are most commonly associated with drugs, which can be beneficial to bacteria as well as humans. The inhibitory

compounds produced by bacteria have a broad spectrum of activity and can have antifungal, antibacterial, anticancer, antiparasitic, and immunosuppressive activities.

In the 1990s and early 2000s, large pharmaceutical companies began to back out of natural product discovery programs due to rediscovery of previously isolated compounds. The process of discovering novel compounds and getting these compounds to market is a laborious process (Walsh and Wencewickz 2014), and the pharmaceutical companies believed that natural product discovery was no longer as beneficial as it once was (Bérdy 2012). Companies turned to high throughput screening (HTS) and combinatorial chemistry for novel drug discovery programs because these methods allowed for large libraries of synthetic compounds to be screened in a short period of time. These methods of discovering novel drugs were deemed more efficient than the traditional methods (Dias et al. 2012). The advantages of screening synthetic compounds is that the synthetic libraries can be designed for specific novel drug targets, but researchers ran into the problem that the synthesized drug would target a particular area of the cell and be a potent inhibitor, but the new compound could not cross the cell wall or would be effluxed out of the cell once it entered, and thus could not reach its target (Bérdy 2012). Although combinatorial chemistry hasn't been as successful in drug discovery as anticipated, the access to variations of natural products is beneficial compared to random chemical libraries (Bérdy 2012) and shouldn't be abandoned.

With an antibiotic resistance crisis on the rise, there has been a renewed interest in returning to natural product discovery. Data show that nature-derived drugs still account for one third (38%) of the 547 recently approved drugs as of 2013 (Partridge et al. 2015). This number accounted for non-molecular natural products and biochemical natural

products that were approved for various targets such as antibacterial and antifungal. Since the discovery of the first antibiotic, many new techniques and technologies have been developed that could aid in accessing these natural products more easily. These new areas of investigation include sociomicrobiology, nanotechnology, bioinformatics in conjunction with mass spectrometry (MS), nuclear magnetic resonance (NMR) and crystallography (Bérdy 2012). Sociomicrobiology is a relatively new area of focus within microbiology, and it investigates group-behaviors of microbes (Parsek and Greenberg 2005). By examining groups of bacteria instead of the individual, researchers can try to attack the group instead of the individual with novel drugs. Nanotechnology is taking antibiotic resistance research to a new level. For example, techniques are being developed that use polymer-based nanoparticles to supplement antibiotics by destroying bacteria protective membranes. These nanoparticles biodegrade safely in the body (Singh et al. 2014).

With the advances in genomic sequencing and bioinformatics, genome mining has become a useful tool to identify the biosynthetic pathways used to synthesize small molecules within microbes. Some of these biosynthetic pathways have been identified and have lead to novel metabolites (Bérdy 2012) using sophisticated bioinformatics analysis programs. For example, antiSMASH (Antibiotics and Secondary Metabolites Analysis Shell) is a bioinformatics program that is capable of identifying biosynthetic loci covering a whole range of known secondary metabolite compound classes such as polyketides, non-ribosomal peptides, bacteriocins, and siderophores. (Medema et al. 2011). Orphan genes, which are genes that are biosynthetic gene clusters for which the corresponding metabolite is still undiscovered (Chiang et al. 2011) are being examined in new drug discovery programs. A large number of orphan gene clusters have been identified through genomic

sequencing (Gross et al. 2007), and although the origin of such genes is unknown, they are thought to be species specific. These gene clusters represent a rich source of unexplored bioactive compounds (Gross et al. 2007). The challenge for researchers is identifying the biosynthetic compounds that are produced from orphan genes (Chiang et al. 2011). Also at the genomic level, silent gene clusters are being examined. These are gene clusters that are not expressed or are expressed at very low levels and their natural products are difficult to detect (Chiang et al. 2011). Researchers are looking for ways to awaken these silent gene clusters (Chiang et al. 2011; Seyedsayamdost 2014).

Another reason to continue looking into natural products is because the majority of the world's biodiversity remains unexplored. There are millions of bacteria in just one gram of soil (Bérdy 2012), and because of the ease of access to soil samples, this is mostly what pharmaceutical companies screened in natural product discovery programs. Soil research has led to some discoveries of natural products that we continue to use in the clinic today, and there are still a large number of soil dwelling microorganisms that have yet to be identified and screened for natural products. Recent research in genomic sequencing has shown that researchers are accessing, at best, 10% of the small molecules that are available in bacteria alone (Seyedsayamdost 2014). This is supported by information showing that the majority of biosynthetic gene clusters that are responsible for small molecule production remain inactive or silent. Finding ways to get to the other 90% of small molecules is important. Natural products are commonly found in extreme environments such as marine niches and ecological niches such as bacterial symbionts of fungi and plants, and can lead to novel drug discoveries (Fischbach and Walsh 2009; Bérdy 2012). There are currently five main drug targets for the antibiotics used today, and the

challenge is to identify drugs that have novel drug targets, and this might be found in one of these silent or inactive gene clusters (Figure 1) (Walsh and Wencewickz 2014).

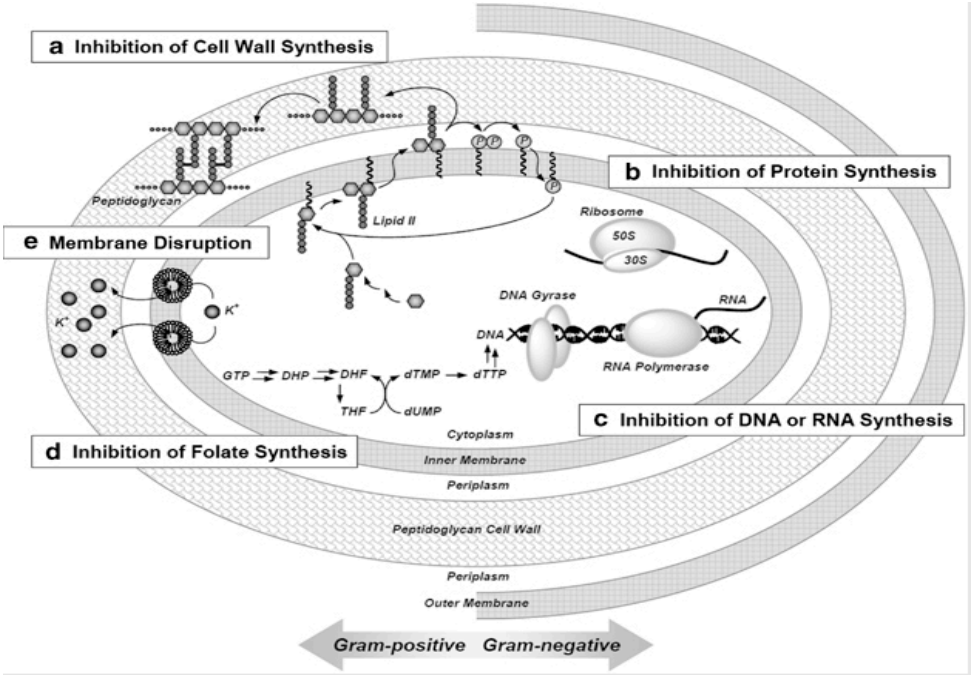


Figure 1: The five major clinically validated antibiotic targets/pathways (Walsh and Wencewickz 2014)

Polyketide Synthases (PKS)

Polyketides are a large chemically diverse class of natural products (Cane and Walsh 1999). Polyketides are found in bacteria, plants, and fungi. Erythromycin (antibacterial), tetracycline (antibacterial), and rampamycin (immunosuppressant) are three examples of compounds produced by polyketide synthases (PKS), a family of multi-domain enzymes (Ridley and Khosla 2007).

Bacterial PKS have been and continue to be essential sources of chemical diversity for drug discovery and development. The complex secondary metabolites that are products

of PKS have impacted all therapeutic areas leading to clinical agents. The explosion of microbial genomic sequencing over the past fifteen years has shown that many organisms encode a wealth of PKS and non-ribosomal peptide synthetases (NRPS), a class of peptide secondary metabolites similar to polyketides (Boddy 2013). PKS assemble their compounds by a biosynthetic assembly line process. The assembly line contains modular synthases (Smith et al. 2015), and the number of modules required for synthesis of the compound often corresponds to the number of building blocks that are incorporated into the final natural product (Haynes and Challis 2007).

There are three types of bacterial PKS that are known to date. These are type I, type II and type III PKS. Type I PKS are made of multi-enzyme domains. Each module has a distinct, noniteratively acting activity that is responsible for the catalysis of one cycle of the polyketide chain elongation (Cheng et al. 2003). Type II PKS are multi-enzyme complexes that carry out a single set of activities. Type II PKS consists minimally of the β-ketoacyl synthase (KS) α and β subunits, and an acyl carrier protein (ACP). Type III PKS are also known as chalcone synthase-like PKS (Cheng et al. 2003). These are predominately in plants, and are the newest to be characterized. They function as condensing enzymes that lack ACP and act directly on acyl-CoA substrates (Cheng et al. 2003).

Among the type I PKS, which are commonly found in bacteria, there is a minimal module containing β-KS, acyltransferase (AT) and ACP. These modules select, activate and catalyze a condensation between extension and the growing polyketide. This generates a β-ketoacyl-S-ACP intermediate. There are optional domains that are found between AT and ACP, which carry out a variable set of reductive modifications of the β-keto group before the next round of chain extension. The order of modules the PKS enzymes dictates the

sequence of biosynthetic events, and the variation of the domains allows for structural diversity among the PKS products (Cheng et al. 2003) (Figure 2).

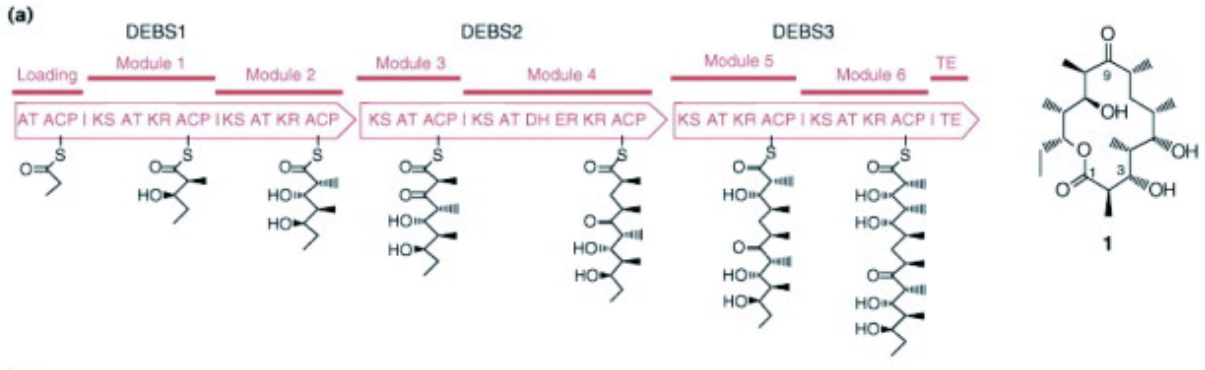


Figure 2: Modular polyketide synthase. A modular PKS that consists minimally of KS, AT, and ACP. With each module, the compound is extended to a final product (1) of 6-deoxyerythronide B (DEBS). KS, ketroreductase; DH, dehydratase; ER, enoylreductase; TE, thioesterase (Tang et al. 2000)

Because PKS produce secondary metabolites, identifying the PKS gene cluster within bacterial genomes is important. If a PKS can be found and characterized in a genome, bioinformatic analysis can be used to potentially identify the composition and structure of the polyketide that the PKS is producing. This allows for an alternative method to discover the identity of a natural product.

The Genus Rhodococcus

The genus *Rhodococcus* is a member of the phylum *Actinobacteria*. Zopf first described it in 1891 to accommodate two species of red pigment-producing bacteria. The genus name as it is currently being used is based on the work of Tsukamura and

Goodfellow and Alderson (Tindall 2014). Currently, there are over 50 species that are classified under the genus *Rhodococcus* (Euzéby 2015).

Members of the genus *Rhodococcus* are described as being aerobic, non-motile, Gram-positive, mycolate-containing, nocardioform actinomycetes. Their characterization as norcardioform refers to their ability to form hyphae that then break into rods or cocci (Bell et al. 1998). They are also known for being GC-rich. GC-rich describes the percentage of nitrogenous bases of DNA in a genome that are guanine and cysteine. *Rhodococcus* are primarily soil-dwelling bacteria, but they have also been found in extreme environments such as the deep sea, Alpine soils, Antarctica, herbivore dung, freshwater and contaminated environments (Bell et al. 1998). *Rhodococcus* have varying shades of pigmentation, and their colony morphology can be rough, smooth, and mucoid in appearance depending on the species (Figure 3) (Bell et al. 1998). The genus has not been characterized well, and overall there are still many questions that remain unanswered.

Figure 3: Variability of pigmentation and morphology in the genus *Rhodococcus*. Rhodococci grow with different pigmentations as well as texture of the colonies (Borisova 2011)

Rhodococci are not known for their pathogenic abilities, but there are a few species within the genus that can be infectious to plants, animals and immunocompromised individuals. These are *R. bronchialis,* which is a strain that has been linked to individuals with cavitary pulmonary tuberculosis and bronchiectasis (Goodfellow 1998). *R. fascians* is another bacterium that is known to infect plants (Goodfellow 1998). This species causes what is called leafy gall disease among plants, and it commonly infects tobacco plants. The most well known infectious species within *Rhodococcus* is *R. equi*. This species infects domesticated animals such as horses and goats, but it can also be pathogenic to immune suppressed patients (Goodfellow 1998). The pathogen causes pneumonia in foals, and the infection in humans resembles pulmonary tuberculosis (Letek et al. 2010).

Industrial Importance

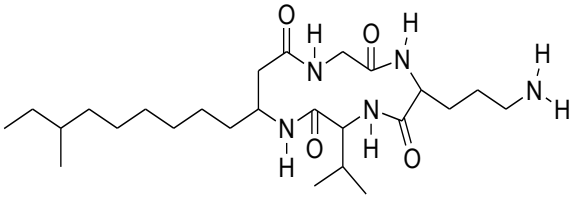
Although the genus *Rhodococcus* is not well characterized, its industrial importance is well known and has been useful for many decades. Diverse oxygenase enzymes contribute to their ability to degrade a broad range of compounds including organic pollutants (Iwabuchi et al. 2002). The production of acrylamide from acrylonitrile by *Rhodococcus rhodochrous* J1 is the most commercially successful chemical biotransformation by a microbe (Kobayashi and Shimizu 1998). Rhodococci have an incredible stress tolerance due to characteristics in their gram-positive cell wall and genetic responses in a stress induced environment (LeBlanc et al. 2008). For example, in *Rhodococcus opacus*, it has been discovered that in a desiccated environment the bacteria produce an extracellular polysaccharide (EPS) slime along with three osmolytes, trehalose, ectoine and hydroxyectoine (Alvarez et al. 2004). Cells also reduced respiratory production

in the stressed environment and utilized fatty acid reserves for energy and biosynthetic precursors (Alvarez et al. 2004). This in combination with their catabolic versatility has led to their wide use in industry.

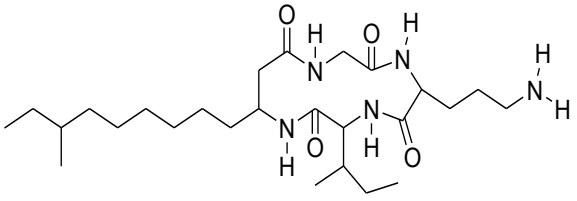
Secondary Metabolites Derived from *Rhodococcus*

Even though little is known about *Rhodococcus,* there have been a few antibiotics that have been discovered from the *Rhodococcus* species. The first report of an antimicrobial compound was from *Rhodococcus* sp. Mer-N1033 in 1999. Chiba, et al. discovered 5 novel cyclic tetrapeptides that showed antifungal activity against *Candida albicans* and *Cryptococcus neoforms* (Chiba et al. 1999). These 5 novel compounds were named rhodopeptin C1, C2, C3, C4 and B5. *Rhodococcus* sp. Mer-N1033 was isolated from a soil sample taken from Mt. Hayachine, Iwate Prefecture, Japan. The rhodopeptins were isolated as colorless solids or white powders, and they were soluble in acetic acid, dimethyl sulfoxide (DMSO), methanol and slightly soluble in water. The five rhodopeptin compounds were insoluble in chloroform and ethyl acetate. Each rhodopeptin was composed of 3 α-amino acids and 1 lipophilic-β-amino acid (Chiba et al. 1999). The structures of the five rhodopeptins are shown in figure 4.

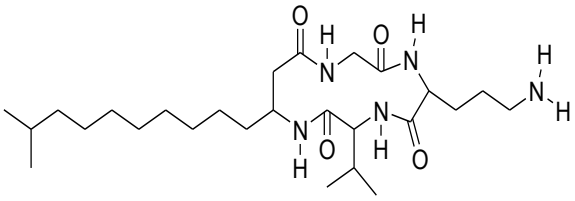
Rhodopeptin C1



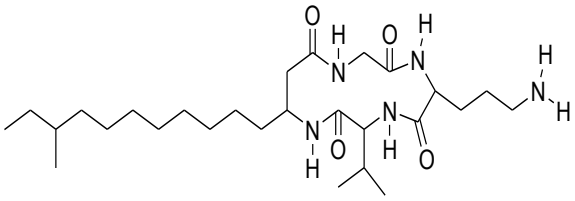
Rhodopeptin C2



Rhodopeptin C3



Rhodopeptin C4



Rhodopeptin B5

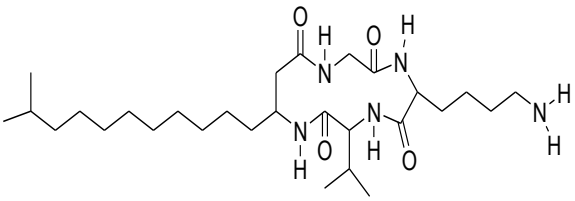


Figure 4: The structure of rhodopeptin C1, C2, C3, C4, and B5

Another antimicrobial compound was isolated from *Rhodococcus jostii* K01-B0171 in 2007 by Iwatsuki, et al. This strain of *R. jostii* was isolated from a soil sample taken in Yunnan, China. This strain was found to produce two antibacterial compounds, and they were named lariatin A and lariatin B (Iwatsuki et al. 2007). Both of the lariatin compounds inhibited *Mycobacterium smegmatis*, and lariatin A also inhibited *Mycobacterium tuberculosis*. It was determined that the compounds are cyclic peptides with a lasso structure. Each compound contained 18 – 20 amino acids with an internal bond between the α-amino group of Gly1 and the γ-carbonyl group of GluB (Figure 5). Lariatin A and B were isolated as pale yellow powders with a molecular weight (MW) of 2050 and 2204,

respectively. Both compounds were soluble in water, DMSO, and methanol and insoluble in ethyl acetate and chloroform (Iwatsuki et al. 2007).

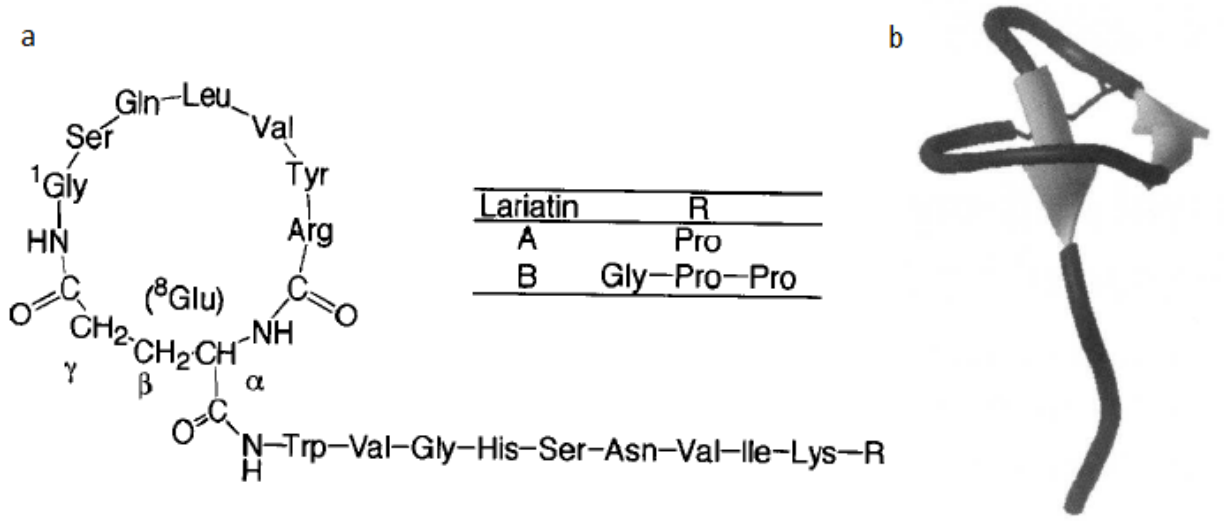


Figure 5: (a) Structure of lariatins A and B. (b) The lasso structure of lariatin A (Iwatsuki et al. 2007)

Another research team, Kitagawa and Tamura, in 2008 discovered fourteen *Rhodococcus erythropolis* strains and one *Rhodococcus globerulus* strain that inhibited at least one of their indicator strains in the study. This massive screen took 80 *Rhodococcus* strains from Japanese and German collections and screened them against *Escherichia coli*, some species of *Pseudomonas, Sinorhizobium, Streptomyces, Corynebacterium, Arthrobacter,* and *Rhodococcus*. After the initial screen, the 15 producing strains of *Rhodococcus* that had been identified were screened again against a different set of 52 bacterial strains to determine the spectrum of activity of each inhibitory compound. The compounds produced by the 15 strains of Rhodocci inhibited many of the Gram-positive indicator strains, but most Gram-negative indicator strains were resistant to the compounds. The inhibitory

compound producers were classified into three categories based on their spectrum of activity. Group 1 consisted of 5 *Rhodococcus* strains that produced antibiotics with activity against a broad spectrum of Gram-positive bacteria. Group 2 contained 3 strains that produced antibiotics against mostly *Rhodococcus* and some other Gram-positive bacteria. Group 3 was 7 strains that exhibited antibiotic activity against only other *R. erythropolis* strains (Kitagawa and Tamura 2008a).

A second publication from Kitagawa and Tamura described how the team continued to work on the structure and characteristics of one of the antibiotic-producing strains in group 1. This strain was *R. erythropolis* JCM6824. The antibiotic produced from this strain was a quinolone structure, aurachin RE (Figure 6), that had a similar structure to aurachin C. Aurachin C is an antibiotic derived from *Stigmatella aurantiaca* Sga15. Both of these antibiotics were found to inhibit the growth of a wide range of Gram positive organisms, but aurachin RE proved to have a much stronger antimicrobial activity. Aurachin RE was isolated as a gray-brown powder, it had a MW of 395, and it was observed to be soluble in ethanol, methanol, ethyl acetate, acetonitrile and DMSO (Kitagawa and Tamura 2008b).

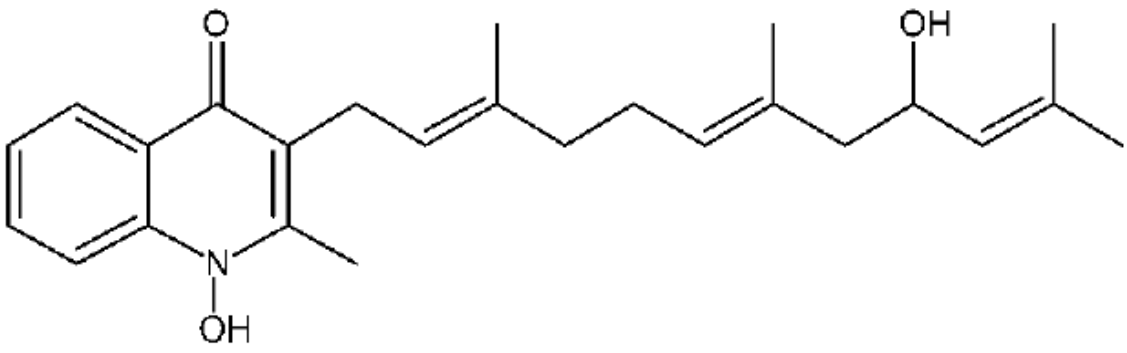


Figure 6: The structure of aurachin RE (Kitagawa and Tamura 2008b)

Finally, a different approach was taken in novel drug discovery, and in 2008, Kurosawa, et al. tested the ability of two organisms, *Streptomyces padanus* and *Rhodococcus facians* to undergo horizontal gene transfer. *S. padanus* is an actinomycete that produces the antibiotic actinomycin. *R. facians* does not produce any inhibitory compounds. After researchers co-cultured the two organisms, a new strain of *R. facians* was recovered and named 307CO. This new strain was found to have a piece of DNA that belonged to *Streptomyces*, and the transferred gene allowed this new *Rhodococcus* strain 307CO to produce two new antimicrobial compounds independently. Researchers named these antibiotics rhodostreptomycin A and rhodostreptomycin B. With further characterization, it was discovered that the two compounds are isomers of a new class of aminoglycosides. It was reported that the structure for the rhodostreptomycin compounds (Figure 7) were distinctly different from the actinomycin compound produced by *Streptomyces padanus.* Rhodostreptomycins inhibited *Escherichia coli, Staphylococcus aureus, Bacillus subtillis,* and *Helicobacter pylori* (Kurosawa et al. 2008).

A **B**

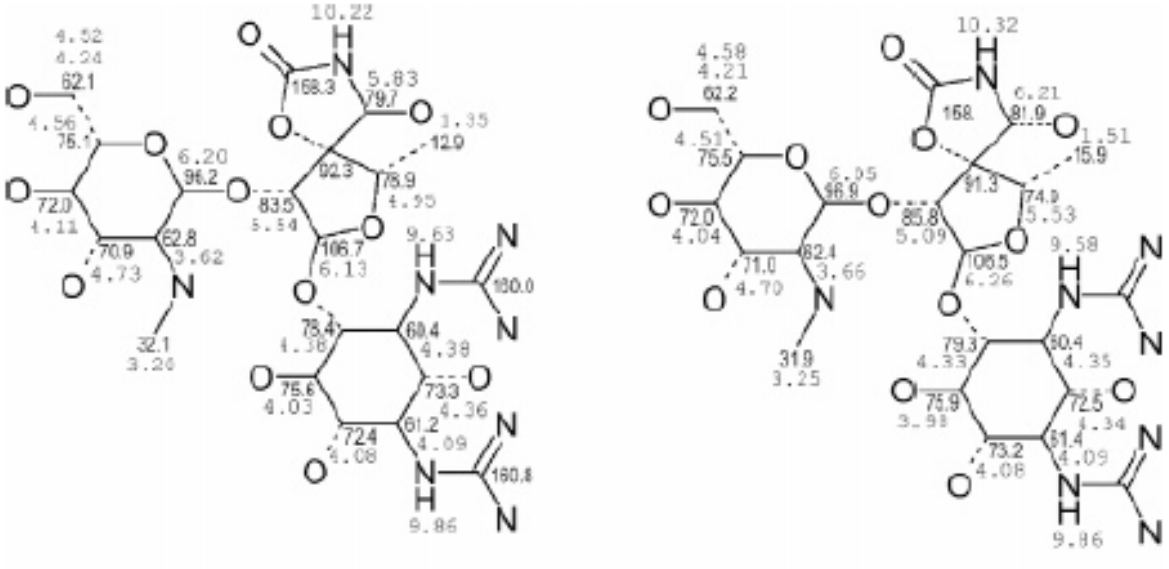


Figure 7: Structure of (A) rhodostreptomycin A and (B) rhodostreptomycin B (Kurosawa et al. 2008)

Several *Rhodococcus* genomes have been sequenced to identify catabolic enzymes, and when reviewing genomic data, many polyketide synthase and non-ribosomal peptide synthetase gene clusters were identified in each of the genomes. Of the fourteen *Rhodococcus* genomes listed in table 2, all of the bacterial genomes have multiple PKS and NRPS regions. The small molecules produced by the various PKS and NRPS genes listed in table 2 are not known. Only the biosynthetic gene cluster required to synthesize the lariatins have thus far been reported (Inokoshi et al. 2012). With these data along with the genus's close relation to *Streptomyces*, the largest bacterial producer of clinically used drugs, *Rhodococcus* should be looked at more closely for natural products in novel drug discovery (Table 2).

Table 2: Proposed biosynthetic gene clusters (genome annotation) in *Rhodococcus*

***Rhodococcus* Genome**	**Genome Size**	**Polyketide Synthase**	**NRPS[a]**
***R. erythropolis* PR4** (Sekine et al. 2009)	6.9 Mb[b]	3	11
***R. ruber* BK** (Bala et al. 2013b)	6.13 Mb	2	>11
R. pyridinivorans (Dueholme et al. 2014)	5.24 Mb	1	>11
***R. equi* 103S** (Letek et al. 2010)	5.04 Mb	1	11
***R. opacus* B4** (Pathak et al. 2013)	8.83 Mb	3	>21
***R. triatomae* BKS 15-14** (Kumar et al. 2013)	5.82 Mb	1	>7
***R. rhodnii* LMG5362** (Pachebat et al. 2013)	4.38 Mb	3	>4
***R. wratislaviensis* IFP2016** (Auffret et al. 2009)	9.69 Mb	>2	>7
***R. jostii* RHA1** (McLeod et al. 2006)	9.7 Mb	>5	>20
***R. qingshengii* BKS 20-40** (Bala et al. 2013a)	6.6 Mb	2	>7
***R. imtechensis* RKJ300** (Vikram et al. 2012)	8.23 Mb	5	>17
***R.* sp. AW25M09** (Hjerde et al. 2013)	5.64 Mb	>4	>5
***R.* sp. P14** (Zhang et al. 2012)	5.67 Mb	>1	>4
***R.* sp. JVH1** (Brooks and Van Hamm 2012)	9.18 Mb	2	>7

[a] NRPS = nonribosomal peptide synthetase
[b] Mb = Megabase pairs
> means at least but probably greater than this number

Rhodococcus sp. MTM3W5.2

Rhodococcus sp. MTM3W5.2 is a bacterium discovered in 2011 (Borisova) that produces an inhibitory compound. This bacterium was isolated from a soil sample in Morristown, Tennessee, located between a driveway and a tomato garden of a local residence. The soil sample was processed using M3 (heat-shock) enrichment. The 16s rRNA gene of this bacterium was sequenced, and the strain was most closely related to *Rhodococcus opacus* with 90% similarity. This strain grows as white colonies that are flat on the agar surface, and as the culture ages, a tan pigmentation is observed (Figure 8). The inhibitory compound produced by MTM3W5.2 is active against *Rhodococcus* and other closely related Gram-positive bacteria (Table 3). It was initially discovered that the inhibitory compound is only produced between 15 - 19°C (Borisova 2011). Since, the inhibitory compound has been produced at temperatures as high as 22°C.

Figure 8: Appearance of *Rhodococcus* sp. MTM3W5.2. The colonies grow flat and spread out widely over the agar surface. MTM3W5.2 initially grows as a white colony and as it ages, the colonies turn tan in color (Borisova 2011)

Table 3: The sensitivity of organisms to the inhibitory compound produced by strain MTM3W5.2 (Borisova 2011)

Organism	Gram reaction	Sensitivity to compound	Diameter of zone (mm)
Rhodococcus erythropolis IGTS8	+	+	36 – 50
Rhodococcus erythropolis DP-45	+	+	17
Rhodococcus equi 33701	+	+	50
Rhodococcus australis 087200	+	+	39
Rhodococcus jostii RHA1	+	+	35
Rhodococcus rhodochrous ATCC 33279	+	+	25
Rhodococcus ruber 1979/002000	+	+	26
***Rhodococcus* sp. MTM3W5.2**[a]	**+**	**-**	**—**
Microbacterium sp. MTM3Y7	+	+	30
Agromyces sp. BEM3Y1	+	+	20
Gordonia sp. BDHXW1B	+	+	27

a: The producer strain

Current Work

As previously stated, *Rhodococcus* is an underexplored genera of bacteria. Based on genomics that identified multiple PKS and NRPS within each sequenced *Rhodococcus* genome and its close relation to *Streptomyces*, *Rhodococcus* may produce many unknown small molecules (Table 2). This is supported by the four groups of inhibitory compounds already discovered within *Rhodococcus*.

Rhodococcus sp. MTM3W5.2 produces an inhibitory molecule that has a narrow spectrum of antibacterial activity. This is encouraging because the inhibitor could be a

novel structure due to its antibacterial specificity. The information about the gene(s) required to synthesize the inhibitor could provide information about the chemical nature of the inhibitor and provide insight into the structure. For this project, I hypothesized that the pTNR transposon system will create random mutations in the *Rhodococcus* sp. MTM3W5.2 chromosome, and that a non-producing mutant will represent an insertional inactivation of a biosynthetic gene required to synthesize or export the inhibitory compound.

CHAPTER 2

MATERIALS AND METHODS

Bacterial Growth Media

Rich Medium (RM)

This medium was prepared by combining the following components:

1) dH_2O	500 mL
2) Glucose (Dextrose)	5 g
3) Nutrient Broth	4 g
4) Yeast Extract	0.25 g
5) Bacto Agar	7.5 g

The ingredients were mixed together and then autoclaved at $121°C$ for 20 minutes. When making RM agar, the media was cooled to 55°C in a water bath, and once cool, the media was poured into sterile Petri dishes (100 mm x 15 mm polystryene, Fisherbrand) and left on the bench top to cool and solidify overnight. The RM agar plates were stored at 4°C until later use.

When making RM broth, the protocol was the same except that the Bacto agar was left out of the ingredients.

M3 Medium (M3)

This medium was described by Rowbotham and Cross in 1977. It was previously reported by Borisova (2011) that there were challenges with precipitation, thus the medium was made in five separate parts. The individual parts were autoclaved separately and were then mixed together after cooling to avoid any of the components from precipitating out. The five components of the medium are as follows for a total volume of 500 mL:

Solution A	
dH_2O	100 mL
KH_2PO_4	0.233 g
Na_2HPO_4	0.336 g

Solution B	
dH_2O	100 mL
NaCl	0.145 g
KNO_3	0.05 g

Solution C	
dH_2O	100 mL
$CaCO_3$	0.01 g

Solution D	
dH_2O	100 mL
Na propionate	0.10 g

Solution E	
dH_2O	100 mL
Bacto agar	9 g

The pH was adjusted to 7.0 for each solution before being autoclaved. After the solutions were autoclaved, they were cooled to 55°C in a water bath. Once cooled, the solutions were mixed together. Trace elements were made into stock solutions and were added to the cooled M3 medium.

Stock Solution	Concentration	Amount per 500 mL of M3
1) $FeSO_4$	1mg/100mL dH_2O	10 μL
2) ZnCl	2.6g/100mL dH_2O	3.5 μL
3) $MgSO_4$	30g/100mL dH_2O	165 μL
4) $MnSO_4$	10mg/100mL dH_2O	100 μL
5) Thiamine HCL	0.4g/100mL dH_2O	0.6 mL

All of the ingredients were mixed well, and then they were poured into sterile Petri dishes as previously described.

Mueller-Hinton Medium (MH)

MH agar plates were made by adding 19.5 g of Difco™ Mueller-Hinton Agar into 500 mL of dH_2O. The mixture was boiled to mix the ingredients, and then autoclaved. After being autoclaved, the medium was cooled to 55°C in a water bath and poured into sterile Petri dishes as described.

MH broth was prepared by adding 11g of BBL™ Mueller-Hinton Broth to 500 mL of dH_2O. The mixture was boiled and then autoclaved. After the broth was autoclaved, the broth was cooled in a 55°C water bath.

Luria-Bertani Medium (LB)

The medium was prepared by combining the following components for 500 mL of medium:

1) Tryptone	5 g
2) Yeast Extract	2.5 g
3) NaCl	5 g
4) Bacto Agar	7.5 g
5) dH_2O	500 mL

The components were mixed well and then autoclaved. Once autoclaved, the medium was cooled to 55°C in a water bath. Once cool, the medium was poured into sterile Petri dishes.

To make LB broth, the same recipe was followed, except Bacto agar was left out.

Bacterial Strains and Growth Conditions

Rhodococcus sp. MTM3W5.2 is the wild-type strain that produces the inhibitory compound. *Escherichia coli* DH5∝ was used for isolating the pTNR plasmid for transformation into *Rhodococcus* sp. MTM3W5.2. *Rhodococcus* sp. MTM3W5.2 was inoculated in RM and grown at 27°C. MTM3W5.2 mutants were grown on the same medium but at 19°C to allow production of the inhibitory compound which is not produced at temperatures above ~21°C. Kanamycin was added to the medium, when appropriate to apply a selection, to a concentration of 75 μg/mL. *E. coli* DH5∝ was inoculated in LB medium and grown at 37°C. Glycerol stocks were made of the *Rhodococcus* sp. MTM3W5.2 strain and it's non-producing mutants. The stocks were made by adding 500 μL of sterile 50% glycerol and 500 μL of an overnight bacterial broth culture. The glycerol stocks were placed in the -70°C freezer for storage.

Creation of Bacterial Seeds

Bacterial seeds were prepared by inoculating a single *Rhodococcus* MTM3W5.2 colony into 2 mL of RM broth. The inoculated broth was placed in a 27°C water bath with shaking for 18 hours. The bacterial seeds were used to inoculate the 10 mL cultures for a genomic prep and plasmid prep. When making a bacterial seed for a genomic prep, the culture incubated for 36 hours instead of the standard 18 hours. The same protocol was used to make a seed for the indicator strain *R. erythropolis* IGTS8 in the disk diffusion assay.

Plasmid Isolation

An overnight bacterial culture of *E. coli* DH5α with pTNR was grown up in 10 mL of LB broth with the addition of antibiotic (Kan^{50} μg/mL) and incubated at 37°C with shaking for 18 hours. The next day, the 10 mL culture was spun down in an Oakridge tube for 10 minutes. After 10 minutes, the supernatant was discarded, and the cell pellet was washed in 1 mL of STE buffer (0.01 M NaCl, 10 mM Tris pH 8.0, 1.0 mM disodium ethylenediaminetetraacetae (EDTA)). 1 mL of the washed cells were transferred to a 1.5 mL microfuge tube and spun in a micro-centrifuge for 45-50 seconds. The supernatant was aspirated off after spinning, and the cells were re-suspended in 200 μL of solution I (50 mM glucose, 10 mM EDTA, 25 mM Tris-HCL, pH 8.0) plus fresh lysozyme (2 mg/mL) and the suspension was left at room temperature for 10 minutes. 400 μL of freshly made solution II (1N sodium hydroxide (NaOH), 10% sodium dodecyl sulfate (SDS)) was added to the suspension, and the microfuge tube was inverted gently several times to mix well. A clearing of the tube was observed, caution was taken to not expose the cells to solution II for more than 5 minutes. 300 μL of cold solution III (5 M potassium acetate (KOAc), pH 5.0)

was added to the microfuge tube, and the tube was inverted several times. A white flocculant formed in the tube, and the microfuge tube was placed on ice for 10-15 minutes. The tube was spun down for 15 minutes at maximum speed (13,900 revolutions per minute (rpm)). Half of the clear supernatant was transferred into a new 1.5 mL microfuge tube, and the other half of the supernatant was transferred to a different 1.5 mL microfuge tube (2 tubes, 450 μL per tube). 200 μL of phenol and 200 μL of chloroform-isoamyl alcohol (24:1) was added to each tube, mixed by inverting the tube several times, and spun for 5 minutes in a micro-centrifuge. After spinning, the top aqueous layer was collected for each tube and transferred to a new 1.5 mL microfuge tube. A chloroform-isoamyl alcohol (24:1) extraction was done twice by adding 400 μL of chloroform-isoamyl alcohol (24:1) to the 1.5 mL microfuge tube and spinning the tube for 1 minute each time keeping the top aqueous layer. The final top layer was collected into a new 1.5 mL microfuge tube, and 1 mL of cold 100% ethanol was added to the tube, mixed, and it was iced for 5 minutes. After 5 minutes, the tube was spun at maximum speed for 10 minutes. The ethanol was removed from the tube using an aspirator, and the plasmid DNA pellet was washed with 400 μL of cold 70% ethanol. The microfuge tube was spun for 2 minutes, and all of the ethanol was removed with aspirator, and the cell pellet was aired dried briefly. After the pellet had dried, the plasmid was re-dissolved in 30 μL of 0.1 TE + RNase (1 mM Tris-HCL, 100 μM EDTA, pH 8.0, 33 μg/mL RNase) and placed in the -20°C freezer for storage.

Solutions for Plasmid Isolation

STE Buffer		Solution I	
2 mL	1 M Tris	5 mL	50 mM Glucose
1 mL	5 M NaCl	2.5 mL	25 mM Tris
0.2 mL	0.5 M EDTA	2 mL	10 mM EDTA
96.8 mL	ddH_2O	Adjust volume to 100 mL with dH_2O	

Solution II		Solution III	
1.4 mL	dH_2O	160 mL	dH_2O
400 μL	NaOH (1N)	98.1 g	KOAc
200 μL	10% SDS	Adjust volume to 200 mL	

Preparation of Electro-Competent Cells

A seed culture of *Rhodococcus* MTM3W5.2 was prepared by inoculating a colony of the MTM3W5.2 culture into 2 mL of RM broth. The seed culture was incubated with shaking at 27°C for 18 hours. A 250 mL flask with 50 mL of RM broth was inoculated with 1 mL of the overnight seed culture. An initial optical density at 600 nanometers (nm) (OD_{600}) was taken. The culture was incubated at 27°C with shaking until an OD_{600} reading of 2-4 was observed. This took approximately 48 hours. A sterile loop of the culture was streaked onto a RM agar plate to test for possible contamination. The agar plate was incubated at 27°C for four days. If there was not any contamination observed, the electro-competent cells could be used. The remaining culture was transferred to a centrifuge bottle (250 mL) and centrifuged in a large rotor (Fiberlite™ F14-6 x 250LE Fixed-Angle Rotor, Thermo Scientific) at 4°C and 6,000 rpm for 7 minutes. The supernatant was discarded, and the cells were re-suspended in 30 mL of ice cold 10% glycerol. The cells were kept on ice while the

re-suspension was being carried out and then iced for 10 minutes. The centrifugation was repeated under the same conditions, and the supernatant was discarded. The cells were re-suspended in 15 mL of ice cold 10% glycerol while on ice, and the cells were transferred to Oakridge centrifuge tubes, iced for 10 minutes, and the centrifugation (Fiberlite™ F14-6 x 50cy Fixed-Angle Rotor, Thermo Scientific) was repeated at 4°C and 6,000 rpm, for 7 minutes. The supernatant was discarded, and the cells were re-suspended in 600 μL of ice cold 10% glycerol. It was important to avoid generating bubbles during this step. Finally, 120 μL aliquots of the re-suspended cells were transferred into small 0.5 mL microfuge tubes, and the electro-competent cells were stored in the -70°C freezer for long-term storage.

The same protocol was followed for electro-competent *E. coli* DH5α cells, except the incubation temperature was 37°C. To reach the optimal OD_{600} of 2 – 4 for *E. coli*, this was obtained within 4 – 5 hours of incubation.

Transposon Mutagenesis

A transposon mutant library was created by electroporation of the pTNR plasmid into electro-competent MTM3W5.2 cells. MTM3W5.2 cells and pTNR DNA was placed on ice for 5 minutes to allow the cells to thaw. 2 μL of pTNR DNA was added to 100 μL of electro-competent MTM3W5.2 cells in a 0.5 mL microfuge tube. The cells were pipetted into a chilled sterile electroporation cuvette (Fisher Scientific, 2 mm width, 400 μL volume) (avoid bubbles) and electroporated at 2,500 volts (V), 600 ohms (Ω), and 10 μF. A pulse time of 3 - 5 was observed. If the pulse time was out of this range, a new microfuge tube of electro-competent cells were prepared for electroporation. 500 μL of RM broth was added

to the cells in the electroporation cuvette, and the cells were transferred to a test tube and incubated at 27°C for 3 hours with shaking. After incubation, the cells were diluted 10 fold and plated on RM Kan^{75}. The plates were incubated at 27°C until colonies appeared. Colonies appeared on day six of incubation and were transferred to RM Kan^{75} wheels. A control was performed by plating *Rhodococcus* sp. MTM3W5.2 onto a RM Kan^{75} plate to ensure that the producer strain would not grow on the kanamycin plate.

Mutant Wheels

After transposon mutant colonies had grown on the RM Kan^{75} plates, each colony was stored on a RM Kan^{75} wheel (Figure 9). Each mutant colony was labeled using the three letters "RMP" then the plate and colony number followed. For example, the sixth colony on plate ten would be labeled as RMP 10.6. These wheels were used to keep a stock of each mutant stored at 4°C while maintaining a selection for only the cells containing the pTNR transposon using kanamycin (75 μg/mL).

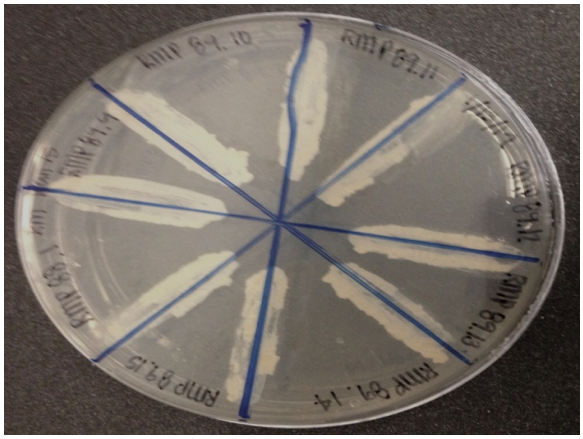


Figure 9: MTM3W5.2 mutant wheels. Each wheel had eight mutants streaked onto the RM Kan^{75} plate. The identification for the mutant was to the left of the growing mutant. These plates were made to keep a stock of the MTM3W5.2 mutants

Preparation of Agar Extracts from RM Plates

This agar extraction method was based on Carr 2012. Each mutant was streaked with a sterile cotton swab from a mutant wheel to make a lawn of growth onto a RM plate that did not have antibiotic added. The RM lawn was incubated at 19°C for two weeks. After two weeks, the plates were cut into 1 cm by 1 cm cubes with a sterile scalpel and placed into a 250 mL beaker without removing the cells from the agar (1 mutant per beaker). 45 mL of ethyl acetate was placed in the beaker with the agar chips, and was slowly shaken to assure that all of the agar chips were exposed to the ethyl acetate (ethyl acetate. HPLC grade, Fisher Scientific). Each beaker was then covered with Parafilm® and left to soak overnight under a hood. After 18 - 24 hours, the liquid extract was decanted into a clean beaker leaving the agar chips behind and left under the hood to dry down. This took another 18 - 24 hours. Once the extract had dried down (Figure 10), 1 mL of methanol (methanol, HPLC grade, Fisher Scientific) was added to the beaker, and a metal spatula was used to scratch the bottom of the beaker to re-dissolve the dried residue. Each extract was then placed in a 1.5 mL microfuge tube for storage at 4°C.

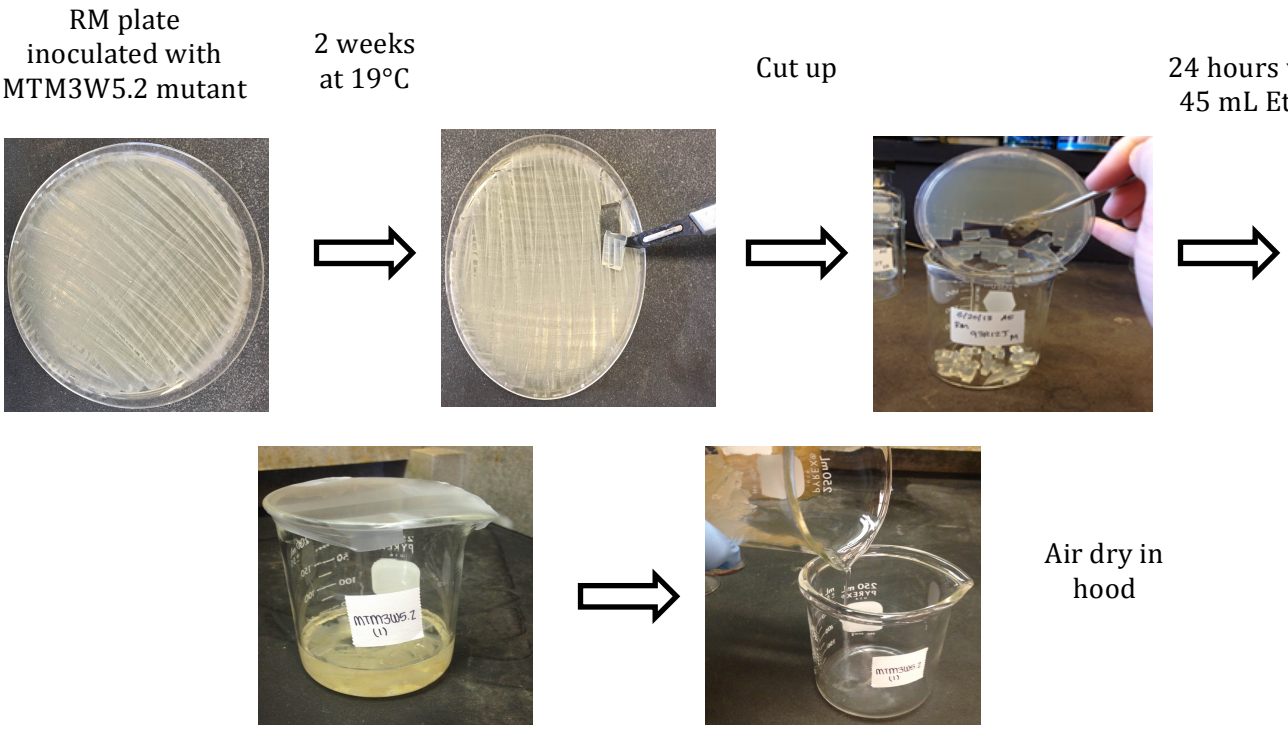


Figure 10: Agar extraction method. Each mutant was grown for 2 weeks at 19°C on a RM agar plate. After two weeks, the plate was cut up into small cubes and placed into a 250 mL beaker with 45 mL of ethyl acetate. The agar chips soaked in the ethyl acetate for 18 - 24 hours and then were decanted into a new 250 mL beaker and allowed to dry down. Once dried down, 1 mL of methanol was used to collect the extract

Disk Diffusion Assay

To detect the presence of an inhibitor molecule, a disk diffusion assay was used to detect growth inhibition of a sensitive indicator bacterium. Paper discs were made from thick Whatman blotting paper GB004 (15 x 15 cm) and punched using a standard hole-puncher. The disks were placed in a glass Petri dish and autoclaved for sterilization. Once autoclaved, the paper disks were labeled using a lead pencil and 50 μL of each agar extract was soaked onto the disk in 25 μL increments. Prior to adding the agar extract to the paper disks, each extract was centrifuged for 2 minutes to spin down any of the cells that had washed off of the agar chips. The paper disks were air-dried. The extracts were tested against *Rhodococcus erythropolis* IGTS8, the sensitive indicator strain. An overnight seed culture of IGTS8 was used along with a sterile cotton swab to inoculate a MH agar plate.

The previously prepared paper disks were placed on the MH plate, and the plates were incubated at 27°C for two days (~48 hours). After two days, the plates were removed from the incubator to observe zones of inhibition against *Rhodococcus erythropolis* IGTS8. When there were mutants that produced no zone or very small zone of inhibition, the mutants were re-plated and underwent another round of extraction for confirmation using the same protocol.

Auxotrophic Mutant Screen

A patch inoculation technique was used to test for auxotrophic mutants. M3 medium was identified as a minimal medium that the wild-type MTM3W5.2 could grow on. After the MTM3W5.2 mutant library was created, each mutant was patch inoculated onto M3 minimal agar plates as well as RM agar plates to identify auxotrophic mutants. After each mutant was patch plated onto M3 and RM, the plates were incubated at 27°C for one week. An auxotrophic mutant was scored as those mutant strains that failed to grow on the M3 minimal agar plates but grew on the RM agar plate.

Genomic DNA Isolation

With a two-day seed, 10 mL of LB broth in a 50 mL flask was inoculated with *Rhodococcus* and covered with a cotton stopper. The flask was incubated ~48 hours at 27°C with shaking. After two days, 120 μL of ampicillin (50 mg/mL) was added to the late log phase culture to give a final concentration of 15 μg/mL. The cells were incubated in ampicillin overnight at 27°C with shaking. After 24 hours, the cells were transferred to

Oakridge tubes and were spun down for 10 minutes. The supernatant was discarded, and the cell pellet was washed in 3 mL of 1X TE buffer (10 mM Tris-HCL, 1 mM EDTA, pH 8.0) and then centrifuged again for 5-10 minutes. Again, the supernatant was discarded, and the cell pellet was re-suspended in 1 mL of TES buffer (50 mM Tris-HCL, 1 mM EDTA, 25% sucrose, pH 8.0) with the addition of fresh lysozyme. The lysozyme was made in a liquid stock (50 mg/mL) with 1X TE buffer, and 1 mL of stock was added to the cells. The cells were incubated with periodic agitation for 3 hours at 37°C in a water bath. After 3 hours, the cells were split into two 1.5 mL microfuge tubes (500 μL per tube), and 30 μL of 20% SDS plus 10 μL of RNase (33 μg/mL) were added to the cells while the microfuge tubes were gently vortexed. The tubes were inverted several times to mix well, and were then incubated at 55°C for 60 minutes in a water bath. 7.5 μl of proteinase K (5 mg/mL) was added to the microfuge tubes, mixed, and incubated in a water bath at 55°C for 15 minutes. 500 μL of phenol was added to the tubes, then mixed, and were centrifuged for 10 minutes to separate the phases. The top aqueous layer was collected and placed into a new 1.5 mL microfuge tube and placed on ice. 500 μL of chloroform-isoamyl alcohol (24:1) was added to the tubes, mixed, and spun for 5 minutes. The top aqueous layer was collected and added to a new 1.5 mL microfuge tube while on ice. This step was repeated two times. After collection of the final aqueous layer, 1 mL of cold 100% ethanol was slowly added to the top of the aqueous solution. The tube was mixed by inverting the microfuge tube several times until a chromosome clot was formed. The chromosome was spooled onto a pipet tip and was transferred into a new 1.5 mL microfuge tube with 400 μL of 70% ethanol. The tube was spun for 1 minute, and all of the ethanol was removed using an aspirator. The

chromosome was briefly air dried while on ice, and then re-dissolved in 100 μL of 0.1 TE. The chromosomal DNA was placed in the 4°C refrigerator for storage.

Solutions for Genomic Isolation

1X TE Buffer		TES Buffer	
5 mL	1 M Tris-HCl, pH 7.5	9.8 mL	50 mM Tris, pH 8.0
1 mL	0.5 M EDTA, pH 8.0	20 μL	0.5 M EDTA
499 mL	ddH_2O	213 μL	40% Sucrose

Southern Blot Analysis

Digestion of Chromosomal DNA

After genomic DNA was extracted from cells, a restriction digest was set up to digest the DNA for Southern Blot analysis. XhoI or SacI (New England Biolabs) were the restriction enzymes that were used to digest the genomic DNA for Southern analysis. The digestion was setup as follows, and was incubated in a 37°C water bath for 1 hour. After incubation, the digestion was stored at 4°C until further use.

1) H_2O	19 μL
2) Genomic DNA	15 μL
3) NEB Buffer #2	10 μL
4) BSA (100X)	1 μL
5) XhoI/SacI enzyme	5 μL
	50 μL

Preparation of the Agarose Gel

A 0.75% agarose gel of size 11 x 16 cm was poured and allowed to solidify. The gel was placed in 1X TBE in an electrophoresis unit.

Loading

The marker (λ/HindIII) and the chromosomal DNA restriction digests were loaded onto the gel. The electrophoresis unit was started at 126 V until the DNA samples had migrated out of the wells, and then the voltage was reduced to 19 V. The gel was run overnight for 16 hours at 19 V.

Staining

After the electrophoresis was completed, the gel was stained with ethidium bromide (70 μL of ethidium bromide (10 mg/mL stock solution) in 200 mL of H_2O) for 45 minutes. The gel was then de-stained in water for 10 minutes with gentle shaking at room temperature.

Photograph the Gel

A ruler was placed along side the gel so that the 1 cm mark was next to the wells of the gel. A photograph was taken with the ruler. The unused areas of the gel were trimmed away using a scalpel, and the bottom left corner of the gel was cut away to identify orientation of the gel (Figure 11a). The tops of the wells were cut off leaving the bridges to show the position of the wells (Figure 11b). The size of the gel was measured for future purposes, and the gel was then transferred to a plastic dish.

A 0.25M HCl solution was added to the gel, and the plastic dish was shaken slowly for 5 minutes at room temperature. The gel was then rinsed with dH_2O. The DNA in the gel was then denatured by soaking the gel in denaturing buffer (0.5 M NaOH, 1.5 M NaCl) for 30 minutes at room temperature with gentle shaking. The gel was then rinsed with dH_2O. A neutralizing buffer (1.5 M NaCl, 0.5 M Tris-HCl, pH 7.0) was then added to the gel, and the gel was soaked in this buffer with gentle shaking for 30 minutes. After shaking, the gel was rinsed with dH_2O. Finally, the gel was soaked in 20X SSC transfer buffer (3 M NaCl, 0.3 M sodium citrate, pH 7.0) for 30 minutes with gentle shaking.

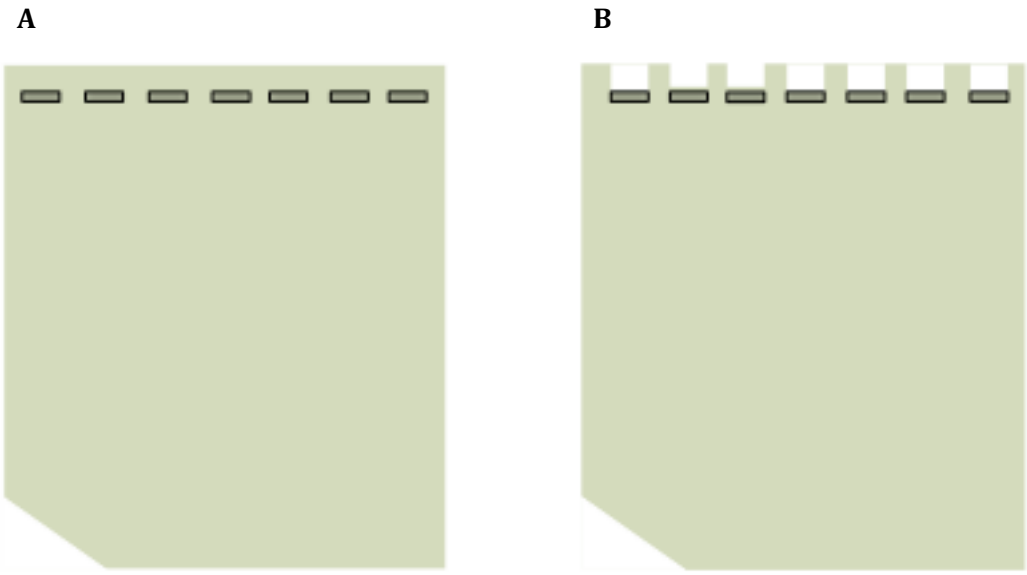


Figure 11: Trimming the gel. (A) The left corner of the agarose gel is cut away to note orientation of the gel. (B) The top of each well is cut away leaving bridges to mark each well

Southern Blot Transfer

A nylon membrane (Nytran^R SPC, 0.45 μm Nylon transfer membrane, 11 x 14 cm) was cut to the size of the gel + 0.5 cm extra on each side. The membrane was then immersed in dH_2O for 30 seconds. The membrane was then soaked in 20X SSC transfer

buffer for 5 minutes. The stack tray of the TurboBlotter™ system (Schleicher & Schuell BioScience TurboBlotter™ Rapid Downward Transfer System) was placed on a flat surface, and 20 sheets of dry, thick blotting paper (Whatman GB004, 15 x 15 cm) were placed in the stack tray. Four sheets of dry, thin blotting paper (Whatman, 3MM Chromatography paper, 20.3 x 25.4 cm) were placed on top of the stack, and 1 sheet of thin blotting paper pre-wet in 20X SSC transfer buffer was placed on top of the stack. After this, the wet nylon membrane was placed on top of the stack. A window of saran wrap was cut to the size of the nylon membrane, and the rest covered the excess blotting paper around the membrane. The agarose gel was then placed on top of the membrane, and there was care taken to make sure that there were no bubbles under the gel. The wells and left corner of the agarose gel were marked onto the nylon membrane using a lead pencil. The top surface of the gel was soaked with 20X SSC transfer buffer. Three sheets of pre-soaked (in 20X SSC transfer buffer) blotting paper were placed on top of the agarose gel (avoiding bubbles). The buffer tray was then attached to the stack tray, and 20X SSC transfer buffer filled the buffer tray. A pre-soaked buffer wick (thin blotting paper, 18 cm x 14 cm) was placed across the stack so that the short dimensions of the wick completely covered the blotting stack, and both ends of the long dimension dipped into the transfer buffer in the buffer tray. A light weight was placed on top of the wick to prevent evaporation of the buffer from the wick. The transfer continued overnight for 18 hours.

Following the transfer, the membrane was gently washed with 2X SSC transfer buffer (300 mM NaCl, 30 mM sodium citrate, pH 7.0) for 5 minutes with gentle shaking. The membrane was then placed on a dry sheet of thin blotting paper to remove any excess 2X SSC transfer buffer. The membrane was then baked at 80°C for 1 hour in a vacuum between

two thin blotting papers. After baking, the nylon membrane was wrapped in aluminum foil and stored at room temperature.

The agarose gel was back-stained with ethidium bromide to ensure that DNA was transferred.

Solutions for Southern Blot

0.25 M HCl (500 mL)

Dilute from concentrated HCl stock (12M)
Add 10.4 mL of 12M HCl stock
Add 498.6 mL dH_2O

Denaturing Buffer (500 mL)

10 g	NaOH
43.83 g	NaCl
400 mL	dH_2O

Neutralizing Buffer (500 mL)

400 mL	dH_2O
30.28 g	Tris
43.83 g	NaCl

Adjust pH to 7.0 with concentrated HCl
Bring final volume to 500 mL

20X SSC Transfer Buffer (500 mL)

400 mL	dH_2O
87.85 g	NaCl
44.1 g	Na Citrate

Adjust pH to 7.0
Bring final volume to 500 mL

Preparation of Labeled Probe DNA

A Thermo Scientific Biotin DNA Labeling kit was used for labeling the pTNR DNA probe. This kit provided everything needed for labeling of DNA. The pTNR DNA was digested with the restriction enzyme HindIII (New England Biolabs). The labeling reaction was initiated by mixing 10 μL of the pTNR DNA, 10 μL decanucleotide in 5X reaction buffer, and 44 μL of nuclease-free water in a 1.5 mL microfuge tube. The mixture was vortexed and

spun down in a micro-centrifuge for 3-5 seconds. The mixture was incubated in a boiling water bath for 10 minutes and cooled on ice. Once cool, the mixture was spun down for 3-5 seconds. 5 μL of Biotin labeling mix (1 mM dGTP, 1 mM dATP, 1 mM dCTP, 0.65 mM dTTP, 0.35 mM Biotin-11-dUTP aqueous solution) was added to the mixture along with 1 μL of Klenow fragment (exo$^-$, 5 u/μL). The microfuge tube was shaken and spun down in a micro-centrifuge for 3-5 seconds. The mixture was then incubated in a water bath at 37°C for 20 hours. After incubation, the reaction was stopped by the addition of 1 μL of EDTA (0.5 M, pH 8.0). The labeled DNA was stored at -20°C until further use.

Hybridization of Southern Transferred DNA

The stored nylon membrane was pre-wet in a small volume of 2X SSC transfer buffer and then placed in a pre-warmed (42°C) glass hybridization tube. The tube was then filled with 12 mL of hybridization solution (50% formamide, 5X SSPE, 5X Denhard's solution, 10% SDS) and placed in the hybridization oven. The membrane was pre-hybridized at 42°C with rotation for 2 hours. Following the pre-hybridization, the hybridization solution was discarded and replaced with 6 mL of fresh hybridization solution that was supplemented with 18 μL of biotin labeled probe DNA. The nylon membrane was hybridized at 42°C for 18 hours with rotation. The hybridization solution was saved in a 50 mL conical tube and stored in the -20°C freezer. It was used up to two times and then discarded.

Hybridization Solutions

Hybridization Solution		20X SSPE	
50 mL	Formamide	87.65 g	NaCl
30 mL	5X SSPE	13.80 g	Na_2HPO_4 H_2O
10 mL	50X Denhardt's Solution	3.7 g	EDTA
10 mL	10% SDS	Adjust pH to 7.4 with NaCl	

50X Denhardt's Solution

1 g	Ficoll 400
1 g	Polyvinylpyrrolidine (PVP)
1 g	Bovine Serum Albumin

Bring to final volume 100 mL with dH_2O

Detection of the DNA

A Chemiluminescent Nucleic Acid Detection Module kit from Thermo Scientific was used for the detection of DNA. This kit provided all solutions necessary for the detection of DNA. The blocking buffer and 4X wash buffer were slowly warmed to 37°C in a water bath. The membrane was then blocked by adding 16 mL of blocking buffer to the membrane and incubating it for 15 minutes with gentle shaking at room temperature. A conjugate/blocking buffer solution was prepared by adding 50 μL of streptavidin-horseradish peroxidase conjugate to 16 mL of blocking buffer. The blocking buffer was decanted after 15 minutes from the membrane, and 16 mL of the conjugate/blocking buffer solution was added to the membrane. This incubated at room temperature for 15 minutes while shaking. The nylon membrane was then transferred to a new plastic container and rinsed briefly with 20 mL of 1X wash solution. The membrane was then washed four times

for 5 minutes each with 1X wash buffer while gently shaking. After the membrane had been washed, it was transferred to a new plastic container, and 30 mL of substrate equilibrium buffer was added to the container to incubate for 5 minutes with gentle shaking. A chemiluminescent working solution was prepared by adding 6 mL of luminol/enhancer solution to 6 mL stable peroxide solution. The membrane was removed from the substrate equilibrium buffer and gently blotted one corner with a paper towel to remove excess buffer. The membrane was then placed in a plastic dish with the nucleic acid side down into a puddle of chemiluminescent substrate working solution. This was incubated in the dark for 5 minutes without shaking. The membrane was then removed from the working solution and a corner of the membrane was blotted onto a paper towel to remove any excess buffer. The moist membrane was wrapped in saran wrap while avoiding bubbles. The blot was imaged using a Protein Simple FluorChem™ M system. An automatic function was used to find the best exposure time for the blot. Most often, the optimal exposure time was 10 minutes.

Recovery of pTNR Insertion Sites from Mutants

Withers et al. (2014) described the procedure for self-ligation of DNA. This protocol was used to self-ligate the mutant DNA fragments together once the DNA had been digested with the restriction enzyme XhoI. The recovery (cloning) of the pTNR insertion sites from the MTM3W5.2 mutants was accomplished by extracting genomic DNA from each mutant as described. Each mutant's DNA was digested with the restriction enzyme XhoI (New England Biolabs). The Thermo Scientific NanoDrop 2000 was used to determine the

concentration of the genomic DNA, and a desired 2 ng of DNA was used for digestion. The setup for the digestion was as follows:

1) Genomic DNA	2 ng
2) H_2O	appropriate volume to achieve 50 μL
3) NEB buffer #2	10 μL
4) XhoI enzyme	5 μL
	50 μL

The digestion was incubated at 37°C for 1 hour. After digestion, 11 μL of digested DNA was mixed with 1.5 μL of 10X T4 DNA ligation buffer (50 mM Tris-HCL, 10 mM MgCL, 10 mM dithiothreitol , pH 7.5, New England Biolabs), 1.5 μL of 100 mM ATP (New England Biolabs) and 1 μL 1X T4 DNA ligase (New England Biolabs) in a 1.5 mL microfuge tube. The mixture was incubated for 2 hours at room temperature (Figure 17). The mixture was heat inactivated in a water bath for 10 minutes at 65°C. 2 μL of the ligation mixture was used to transform electro-competent *E. coli* DH5α cells as described. The transformed cells were plated onto LB Kan^{50} plates and incubated at 37°C for 48 hours until colonies appeared. LB Kan^{35} plates were also used, and the transformed colonies would appear within 24 hours.

Sequencing of pTNR Insertion Sites

The insertion regions of pTNR were sequenced by designing primers that read out from the transposon termini into the gene of interest (Figure 17). The ligated plasmid DNA was extracted from the *E. coli* DH5α cells using a Qiagen Qiaprep Spin Miniprep Kit. In the first round of sequencing, the primers that were used were designed based off of the pTNR

transposon sequence. The primers were (5'-TGAGTGCTTGCG GCAGCGTCTAG-3' and 5'-GATCCTTTGATCTTTTCTACGGGG-3'). The primers (3 μM) and the plasmid DNA (100 ng/ μL) were sent to the Molecular Biology Core Facility at East Tennessee State University for Sequencing. This facility uses a Beckman CEQ 8000 automated DNA sequencer.

Analysis of Sequenced Insertion Sites

The NCBI basic local alignment search tool (BLAST) was used to analyze the sequenced insertion site of each non-producing mutant. After several rounds of sequencing for one mutant, the overlapping sequencing "runs" were combined to form one continuous DNA sequence, and this was analyzed by NCBI blastx for similar protein sequence alignments. This database searches protein databases using a translated nucleotide query. Results were analyzed to look for similarities between the sequencing data and known polyketide synthase regions within *Rhodococcus* and other bacteria.

Scale-Up Production of Antimicrobial Compound

A method was needed to increase the production of the inhibitory compound in order to purify, isolate and identify the compound. The scale-up was done by extracting from a large (150 mm x 15 mm) Petri dishes containing RM agar. A two-day 2 mL seed of MTM3W5.2 was grown at 27°C with shaking. After two days, a sterile cotton swab was used to inoculate the large Petri dish for a lawn of growth. Fifteen large Petri plates were used for one scale-up. The plates were incubated at 19°C for two weeks. If a plate was found to have a single colony contamination, the contaminant was cut out of the agar, and the rest of

the plate was used for extraction. The plates were cut into 1 cm by 1cm square cubes. The agar chips were placed in a 2 L beaker along with 1 L of ethyl acetate. The agar cubes were stirred around to remove any air in the beaker. The beaker was covered with Parafilm® and left to soak for 18 - 24 hours. After 18 - 24 hours, the extract was decanted into a clean 2 L beaker leaving the agar chips behind and left to evaporate under the hood. 1 L of fresh ethyl acetate was added to the 2 L beaker with agar chips, stirred, covered and left to soak for another 18 - 24 hours. After 18 - 24 hours, the extract was decanted into the beaker with the previous days extract and the combined extract was left to sit under the hood to evaporate completely. The large scale-up extract took 4 – 5 days to dry down under the hood. Once the extract had dried down, the dried residue was re-dissolved in 6 mL of methanol by scraping the bottom and sides of the beaker using a metal spatula. The extract was then poured into a 15 mL polystyrene tube and centrifuged at 8,000 RPM for 30 minutes to remove any of the un-dissolved particles. The centrifuged extract was then transferred to a separatory funnel along with the same volume of deionized water (~ 6 mL). This wash with deionized water was repeated three times, and after separation of the organic layer (top layer in separatory funnel), the organic layer was evaporated under the hood overnight. Once evaporated, the dried extract was dissolved with 10 mL of methanol and washed three times with 10 mL of n-pentane (+99%, extra pure, Fisher Scientific). After collecting the methanol (bottom layer), the methanol extract was evaporated under the hood again and the dried residue was dissolved in 1 mL of 100% ethanol. The extract was stored at 4°C until use.

Sephadex LH-20 Column Chromatography

Preparation of the Sephadex LH-20 column was described by Wright (2010). The Sephadex LH-20 column separates compounds based on their hydrophobicity and molecular weight. The column was prepared by suspending 25 g of LH-20 resin in methanol and de-aerating by gently shaking for 20 minutes. After de-aeration, the material was packed into a 1.5 x 50 cm column to a depth of 45 cm. The column was then equilibrated with three bed volumes (1 bed volume = ~ 80 mL) of methanol. After equilibration, 5 mL of the scaled-up extract was loaded onto the resin and allowed to settle for five minutes. After five minutes, the column was exposed to a constant flow of methanol as the mobile phase. Column fractions were collected using a Bio-Rad 2110 fraction collector. 50 fractions were collected with each fraction being 175 drops (2.5 mL) each. After all of the fractions were collected, every other fraction was tested for antimicrobial activity using the disk diffusion assay described previously. Fractions that were identified as having activity were pooled together and placed in a 250 mL beaker and left under a hood to evaporate. Once completely dry, the dried residue was re-dissolved in 7 mL of 90% methanol. This was made by adding 6.3 mL of 100% methanol and then 0.7 mL of dH_2O drop by drop to avoid precipitation. The extract was stored at 4°C until needed for further use.

High Pressure Liquid Chromatography (HPLC)

The next step in the purification process was to run the extract through a BioRad Biologic Duoflow High Pressure Liquid Chromatography Column for better separation of

the compounds that are contained in the MTM3W5.2 extract. The extract was run in 1 mL aliquots using a Waters 7.8 mm x 300 mm Novapak HR C_{18} hydrophobic column as the stationary phase. The mobile phases were de-aerated dH_2O and methanol. The program that was used was designed and described by Borisova (2011). The mobile phase started at 90% methanol and gradually increased the methanol concentration until it reached 100%. The presence of the compounds was detected by ultra violet (UV) light that was set at 254 nm. 35 fractions were collected, and every other fraction was tested for antimicrobial activity to determine the peaks that corresponded to the inhibitory compounds. Fractions that showed activity were compared to the peaks on the chromatogram, and the active fractions were pooled together and stored at 4°C.

CHAPTER 3

RESULTS

Generation of Mutant Strains Using pTNR

Mutations were created in *Rhodococcus* sp. MTM3W5.2 DNA using the plasmid pTNR. The pTNR plasmid was designed specifically for *Rhodococcus* genomes, and it contains the insertion sequence (IS) element 1415. The pTNR plasmid will not replicate when introduced into *Rhodococcus* cells by electroporation. The plasmid was designed so that the genes *istA* and *istB* encode the transposase (Figure 12). This is the protein that is necessary for the movement of the IS element from one site to another, and no other genetic material is encoded in the *istA* and *istB* region of the plasmid. pTNR was also designed such that the orientation of the inverted terminal repeats (IR1 and IR2) are reversed. During transposition therefore, the kanamycin resistance gene (Kan^r) and plasmid origin of replication for *E. coli* (*ori*) that are encoded between IR1 and IR2 are moved into the genome, and the transposase, *istA* and *istB,* remains on the plasmid and is lost. The origin of replication enables the inserted DNA in the *Rhodococcus* genome to be cloned and replicate like a plasmid within *E. coli*, and selection for transposed mutants can be made using kanamycin (Sallam et al. 2006).

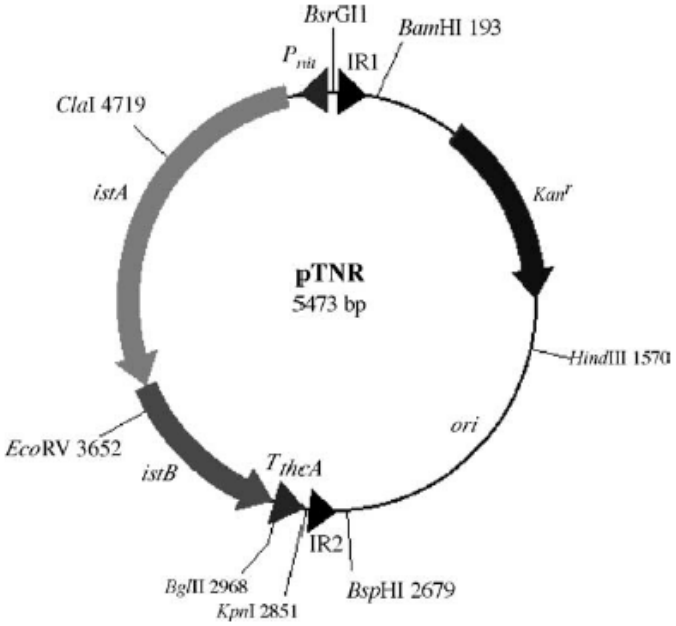


Figure 12: Map of pTNR plasmid. The genes encoded in the pTNR plasmid are *istAB*, genes for transposition from activity from IS*1415*; IR1, inverted repeat 1 upstream of *istA*; IR2, inverted repeat 2 downstream of *istB*; Kan^r, kanamycin resistance; *ori*, origin of replication for *E. coli*; T_{thcA}, transcriptional terminator; P_{nit}, *nit* promoter (Sallam et al. 2006)

The IS*1415* element from the pTNR plasmid acts as an insertional mutagen. This mutagen will insert a piece of DNA, the region of the plasmid between IR1 and IR2 into the *Rhodococcus* genome. The insertion of the DNA can interrupt a gene's reading frame, which inactivates the gene's function. Electrotransformation was used to transform electro-competent *Rhodococcus* sp. MTM3W5.2 cells with the pTNR plasmid (Figure 13). This plasmid was chosen to create the MTM3W5.2 mutant library because it has been used successfully on *Rhodococcus* genomes in previous studies (Sallam et al. 2006), and it usually produces single insertions throughout the genome (1 insertion per mutant).

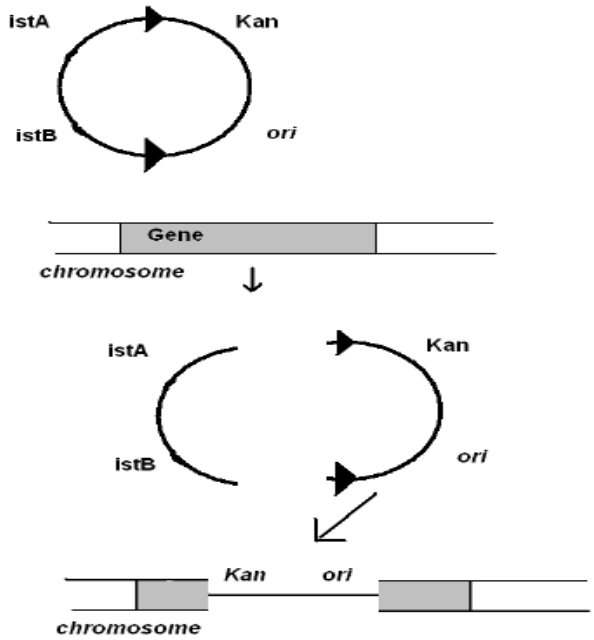


Figure 13: Mutagenesis with IS*1415* element of the pTNR plasmid. The transposed DNA, containing the Kan^r gene and *ori* site, randomly inserts into genomic DNA, interrupting a gene's function (Pratt 2008)

After electroporation, the transformants were selected by plating 50 μL of a 10^{-1} dilution of transformed *Rhodococcus* MTM3W5.2 cells onto RM agar plates that contained 75 μg/mL of kanamycin. The addition of kanamycin selects for transformants that only contain the kanamycin resistance marker of pTNR. This ensures that there were only pTNR mutants being selected for. The agar plates were then incubated for 5 – 6 days at 27°C. A total of 2,306 kanamycin resistant colonies were recovered and screened using an auxotrophic screen and agar extraction method (Table 4).

Auxotrophic Mutant Screen

The presence of an auxotrophic mutant indicates that there are mutants being created in the *Rhodococcus* genome by the pTNR transposon. An auxotrophic mutant is a

mutant that cannot synthesize an essential nutrient such as one of the twenty amino acids when inoculated on a minimal growth medium. An auxotrophic mutant has the ability to grow on a rich medium that contains all of the essential nutrients. A minimal medium only contains basic salts and glucose as a carbon source, but it lacks other essential nutrients and cannot support the growth of an auxotroph. Each of the 2,306 mutants was patch-plated onto M3 minimal medium agar plates and RM agar plates. The mutants that were able to grow on RM agar, but not on the M3 minimal medium were scored as auxotrophic mutants (Figure 14 denoted by the black circle). Of the 2,306 Kan^r colonies, 7 were identified as auxotrophic mutants.

Table 4: Number of pTNR mutants screened

	Mutant Type	
Total Kan^r Colonies Screened	**Auxotrophic**	**Non-producers**
2,306	7	8

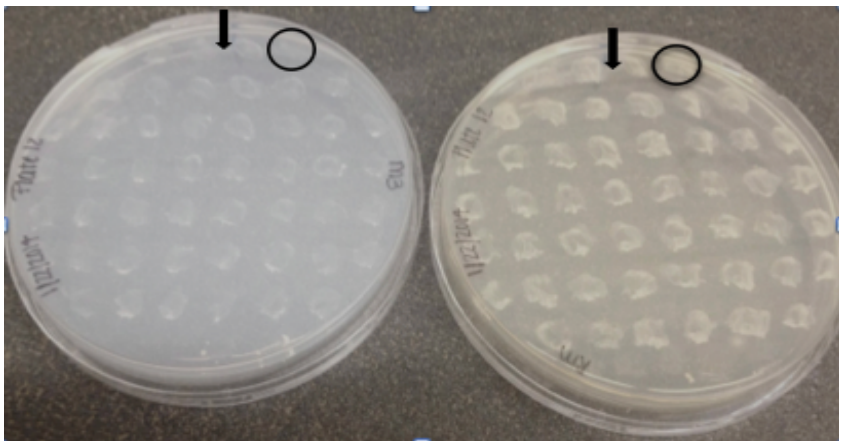


Figure 14: Auxotrophic mutant screen. Mutants that were not able to grow on M3 agar, but grew on the RM agar plate were identified as auxotrophic mutants. As indicated by the circle, there is no growth of the mutant on M3 (left), but there is growth on the RM agar plate (right)

Screening for Non-producing Mutants

The 2,306 Kan^r colonies were screened by an agar extraction method to identify mutant strains that no longer produce the inhibitory compound. The inhibitory compound is not reliably recovered from broth cultures. For example, no compound was recovered from broths supplemented with a binding resin (Borisova 2011). It was also absent from shaking RM broths. In an attempt to replicate Borisova's results, broth cultures were tested with resin and without resin. Both of these two cultures were also tested under the conditions of shaking or stagnant cultures. The broth cultures that contained resin were inconsistent in their disk assay results (some cultures produced the compound, while others did not). Broth cultures that did not contain the binding resin, whether they had been shaking or not, did produce the inhibitor, but there was a wide variability. Overall, stagnant cultures without the binding resin seemed to be the most reliable method amongst broth cultures, but it was determined that agar plate cultures can be used to consistently and reliably detect the inhibitory compound among pTNR mutants.

For the agar-based method, each mutant was plated onto a RM agar plate and incubated at 19°C for two weeks. Previous work discovered that the inhibitory compound is only produced between 15 and 19°C in MTM3W5.2 (Borisova 2011). Work on this project confirmed that the inhibitor could still be produced as high as 22°C. After two weeks, the plates were extracted using ethyl acetate, and the dried residue was re-dissolved with methanol and tested by the disk diffusion assay.

Disk Diffusion Assay

The disk diffusion assay is also known as the Kirby-Bauer method or antibiotic sensitivity test. In this assay, paper disks were soaked with agar extracts prepared from each of the 2,306 kanamycin resistant colonies to detect the presence of the inhibitor compound (Figure 15B). Previous work showed that *Rhodococcus erythropolis* IGTS8 is hypersensitive to the inhibitor compound produced by MTM3W5.2, and thus was chosen as the indicator strain for disk diffusion assay (Borisova 2011). The agar extracts from the producing parental strain, MTM3W5.2, produces ~ 33 mm zone of inhibition (Figure 15A) against the indicator strain, *Rhodococcus erythropolis* IGTS8. pTNR mutants were scored as a non-producing mutant if they had no or a small zone (< 12 mm) of inhibition against the indicator strain (Figure 15C). Of the 2,306 mutants, 8 mutants were identified as non-producing mutants. The 8 non-producers were inoculated onto RM agar plates and extracted a second time using the same method to confirm that a false positive hadn't been discovered. All 8 of the mutants were confirmed to be non-producing mutants.

Rhodococcus erythropolis IGTS8 and *Rhodococcus* N5-59 were used as negative controls. These two strains of *R. erythropolis* do not produce any known inhibitory

compounds, and they were tested with the agar extraction method to ensure that the extraction of agar alone was not causing false zones of inhibition (Figure 16D). A disk that contained only methanol was also tested against *R. erythropolis* IGTS8 to confirm that the methanol alone wasn't inhibiting the indicator strain (Figure 15E). As expected, there was no zone of inhibition for the controls.

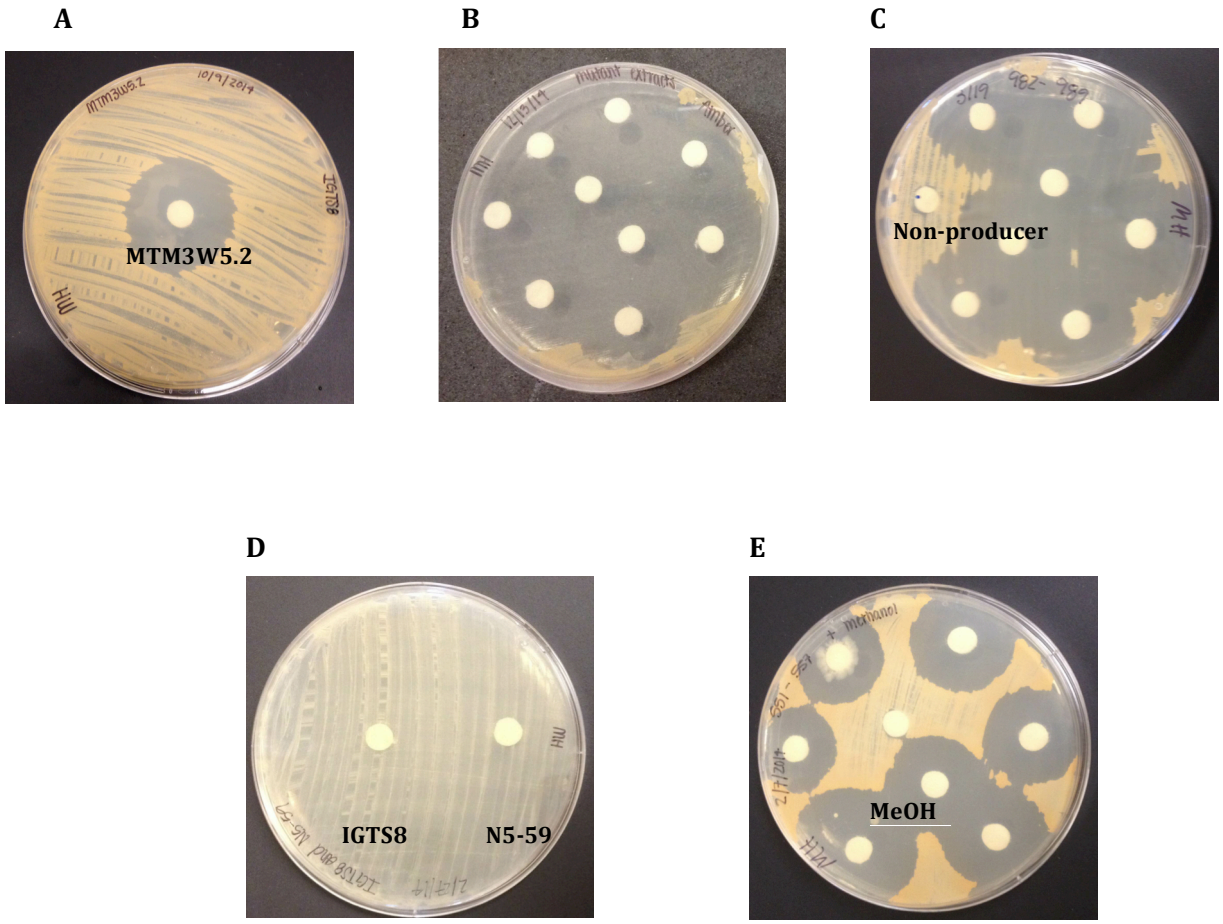


Figure 16: Disk diffusion assay to detect mutant strains that no longer produce the inhibitor molecule. (A) A paper disk, soaked with an agar plate culture extract from the parental strain MTM3W5.2, was placed on an indicator bacterium sensitive to the inhibitor molecule. The zone of inhibition is 33 mm. (B) Paper disks were soaked with agar plate culture extracts prepared from each of 2,306 mutant strains and placed on the indicator bacterium. (C) One of the mutant strains no longer produces the inhibitor compound (paper disk positioned at 9 O'clock). (D) A negative control in which disks are soaked with an extract from a *Rhodococcus* strain that does not produce an inhibitor molecule IGTS8 (left disk) and N5-59 (right disk). (E) A negative control in which one disk is soaked in only methanol, the solvent used to produce the extracts

Southern Blot Analysis of 8 Non-producing Mutants

A Southern hybridization experiment was used to detect transposon insertions in the genomic DNA of *Rhodococcus* species MTM3W5.2. This experiment can confirm that there are single insertions in the MTM3W5.2 genome by the pTNR plasmid, and that the insertions are at different locations within the genome. Figure 16A shows the separation of genomic DNA digested with XhoI from each of the 8 non-producing mutants (after electrophoresis on a 0.75% agarose gel). "L" is the λ DNA marker that is digested with the restriction enzyme HindIII. This is a molecular weight marker that is used to determine the size of the DNA fragments. Lane 1 contains, the genomic DNA of the producer, MTM3W5.2, that was digested with XhoI. This is the negative control for the experiment (does not contain any pTNR DNA). Lane 2 – lane 10 contain mutant strains digested with XhoI and the restriction fragment that hybridized to the pTNR probe (Table 5). Lane C is the positive control, pTNR DNA. One of the mutants did not hybridize to the probe (RMP 71.7 lane 2). This mutant did not hybridized on any of the Southern blots. Further testing should be done to confirm that there is an insertion from the pTNR plasmid. RMP 71.7 did obtain sequencing data, and the initial sequencing primers were designed to attach to the insertion termini of the pTNR insertion region. This suggests that there was an insertion. SacI was also used to digest the genomic DNA, but XhoI was eventually used because it is the same restriction enzyme that is used to recover the insertion regions in each mutant.

Table 5: Size of restriction fragments of the 8 non-producing mutants that hybridize to the pTNR probe

Mutant Strain	**Size of Restriction Fragment Hybridized to pTNR Probe (in kilobase pairs)**
Lane 2: RMP 71.7	No hybridized fragment
Lane 3: RMP 71.3	>23 kb, 23 kb, 9.4 kb
Lane 4: RMP 70.47	23 kb
Lane 5: RMP 82.5	23 kb and 9.4 kb
Lane 6: RMP 46.37	6.5 kb
Lane 7: RMP 46.43	>23kb ,9.4 kb, 4.3 kb
Lane 8: RMP 2.31	6.5 kb
Lane 9: RMP 77.23	6.5 kb
Lane 10: Blank	-----

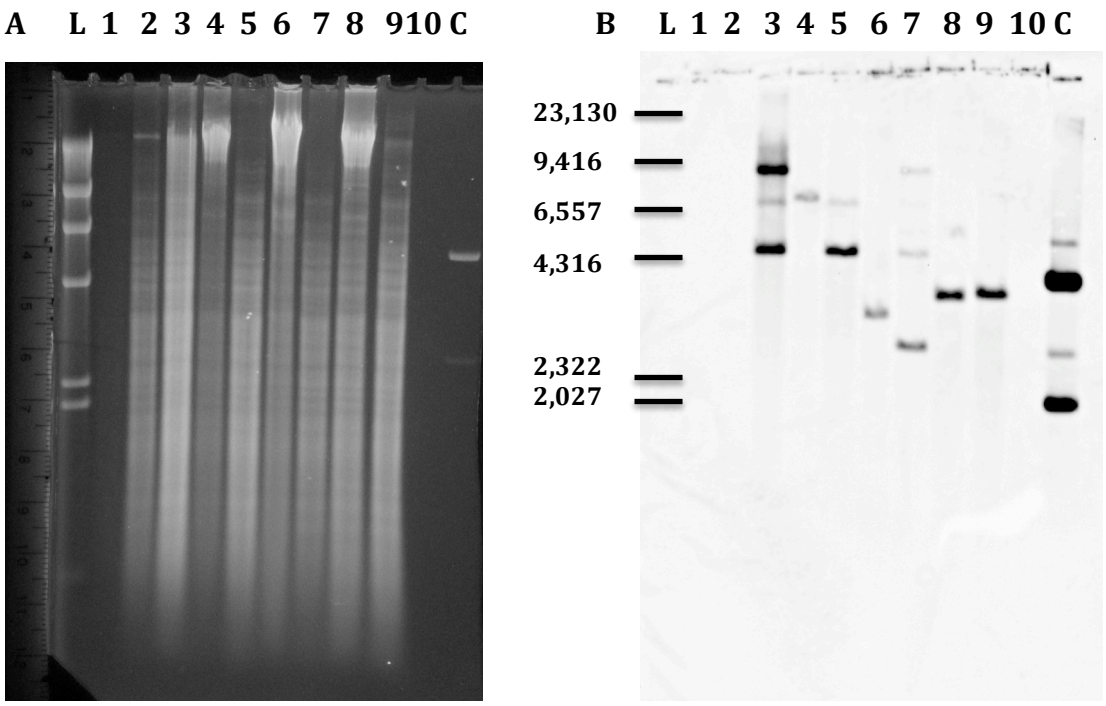


Figure 17: Southern blot analysis of pTNR insertions. (A) Total DNA of each non-producing strain (lanes 2-9) was digested with XhoI and chromosome fragments were separated via agarose gel electrophoresis. Lane L contains lambda DNA size standards, lane 1 contains XhoI digested chromosome from a negative control strain, lane 10 contains no DNA, and lane C contains pTNR plasmid DNA as a positive control. (B) A Southern blot of the agarose gel in (A) was hybridized to pTNR plasmid DNA serving as a probe. Except for lane 2, the non-producing strains contain a hybridizing restriction fragment that varies in size from about 4.3 kb (lane 7) to >23 kb (top band in lane 3). This confirms that the pTNR transposon has inserted into different locations in the chromosome

For each mutant DNA, the XhoI restriction fragment that hybridizes with the pTNR probe is of a different size (Table 5 and Figure 16). This indicates that the pTNR transposon has inserted into different locations in the chromosome of each mutant strain. Some of the mutant strains appear to have a single band that appears on the hybridized blot. This most likely represents a single transposon insertion in the chromosome of these mutant strains. The chromosome from mutant strain RMP 71.3 (lane 3), RMP 82.5 (lane 5), and RMP 46.43 (lane 7) clearly shows multiple hybridizing bands. If the entire DNA was digested, this

could indicate a double transposon insertion. The agarose gel suggests that the DNA was digested. In lanes 4, 6, and 8 on the agarose gel, it looks as if some of the DNA was not digested, but there are only single insertions hybridized on the blot. This may mean that the lanes with multiple signals are there due to double insertions or non-specific binding by the probe to an area in on the chromosome, and that it isn't due to undigested DNA. All transposons have error rates, and it is possible for the pTNR transposon system to insert multiple times in the chromosome.

Recovering pTNR Insertion Site of Mutants

After the 8 non-producing mutants were identified, the gene that was interrupted by the pTNR transposon needed to be recovered by cloning. To recover the pTNR insertion sites, a restriction enzyme, XhoI, was chosen because it will cut within the flanking chromosome but it will not cut within the transposon at the insertion site (Figure 17). This was determined based on the pTNR DNA sequence of the region encoded between IR1 and IR2 (Appendix). After digestion, the excised chromosome pieces were self-ligated together to form a circle (Figure 17). Each ligated circle contains the *ori* site engineered into the pTNR DNA and can function as a plasmid. These cloned plasmids were recovered by transformation of *E. coli*. Some of the plasmids containing the cloned insertion sites are the same size, but several are of different sizes (Figure 18). The size of the ligated plasmid is dependent upon the site where XhoI cut within the *Rhodococcus* chromosome and the transposon. The size of the insertion region is known, and thus can be used to determine the size of the flanking DNA.

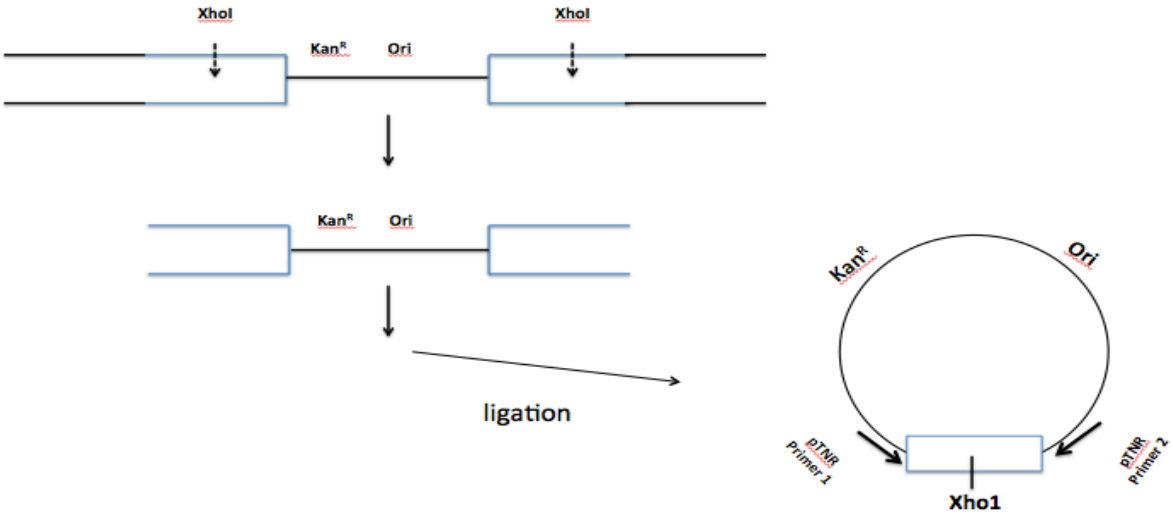


Figure 17: Cloning the insertion site of pTNR mutants. The restriction enzyme XhoI was chosen because it will cut the flanking chromosome DNA but not the transposed pTNR DNA in each mutant. After restriction, the excised chromosome fragment is self-ligated into a circle. Because the transposed DNA contains an ori site, each circle can function as an autonomous plasmid and is recovered by transformation into *E. coli*. The open rectangle is cloned chromosome DNA flanking the insertion site of each mutant. Arrows are primers used to sequence the cloned DNA around the insertion site of each mutant

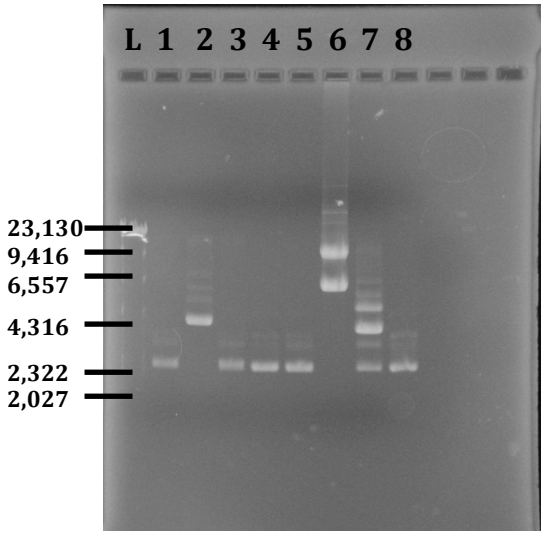


Figure 18: Agarose gel of 8 mutant cloned plasmids. The plasmids were recovered from *E. coli* and ran on a 0.75% agarose gel. L denotes the **λ** ladder that was digested with HindIII. The 8 mutants were in the following order on the gel: lane 1 RMP 70.47, lane 2 RMP 77.23, lane 3 RMP 46.43, lane 4 RMP 71.7, lane 5 RMP 82.5, lane 6 RMP 71.3, lane 7 RMP 2.31, lane 8 RMP 46.37. Several of the cloned plasmids were of similar size. RMP 71.3 was the largest plasmid at ~6.5 kb in size

DNA Sequence Analysis of Cloned Transposon Insertion Sites

For each of the 8 non-producing mutants, DNA was sequenced from the pTNR insertion region contained in the recovered plasmid clones. Obtaining sequencing data for each of these mutants provided information on the identity of the gene that was interrupted by the transposon. After sequencing data was obtained from the ETSU Molecular Biology Core for each mutant, the data were examined using a computer program, MacVector, and then the sequence was transferred to the NCBI Blastx database to search for similarities of translated nucleotide sequence to *Rhodococcus* and other bacteria. Of the 8 mutant clones, one plasmid, RMP 70.47, does not have any sequencing data. Multiple attempts to sequence this mutant were made and all were unsuccessful.

There were two mutants of the 8 that were chosen to study further. One of the mutants of interest, RMP 2.31, has a greater than 55% similarity to known polyketide synthase regions within *Streptomyces* (Figure 19A), and it also shows >45% similarity to polyketide synthase regions in *Rhodococcus* species (Figure 19B).

A

rifamycin polyketide synthase [Streptomyces roseosporus NRRL 11379]
Sequence ID: gb|EWS90296.1| Length: 5445 Number of Matches: 9

Range 1: 1537 to 1754 GenPept Graphics ▼ Next Match ▲ Previous Match

Score	Expect	Method	Identities	Positives	Gaps	Frame
166 bits(419)	3e-67	Compositional matrix adjust.	126/219(58%)	146/219(66%)	7/219(3%)	-2

```
Query  872   AALSEQAARLVDFLGATGDValaeiagallrgraaLDHRAVVLAGAVDDLIRGLTDVAAG  693
             A L+ QA RL   FL    D  L  +A +L    R++L  RAV++A    + + GL  +A G
Sbjct  1537  AGLAGQAGRLAGFLAGRADDELVGVARSLAGSRSSLMDRAVIIASGRGEAVAGLEALARG  1596

Query  692   A--PGL---TGSRVNGRVAFVFPGNRVPSGVGMGRELLVSDSVFATRLAECAAVMDPLTG  528
                PG+    G    GRV FVFPG       VGMG +LL S  VFA RL ECA V+DPLTG
Sbjct  1597  EAVPGVVTAAGESTPGRVVFVFPGQGA-QWVGMGADLLGSSGVFARRLGECARVLDPLTG  1655

Query  527   WSLLDFVAGPQEL-LDRVDVVQPVSFAVMVSLAAVWESCGVTPDVVVGHSQGEIaaacva  351
             WSLL+ V G    + LD VDVVQPVSFAVMVSLAAVWE+CGV PD VVGHSQGEIAAA V
Sbjct  1656  WSLLEVVRGVGGVSLDGVDVVQPVSFAVMVSLAAVWEACGVVPDAVVGHSQGEIAAAVVG  1715

Query  350   galsladaAKVVVARSRVIAQRLAGHGGMVSIALPSERV  234
             G LSL DAA+VVV RSR+I  RL+G GGMVS+A   ERV
Sbjct  1716  GVLSLEDAARVVVERSRLIGARLSGGGGMVSVAAGVERV  1754
```

B

polyketide synthase [Rhodococcus jostii]
Sequence ID: ref|WP_011596689.1| Length: 3527 Number of Matches: 6
▹ See 1 more title(s)

Range 1: 485 to 712 GenPept Graphics ▼ Next Match ▲ Previous Match

Score	Expect	Method	Identities	Positives	Gaps	Frame
125 bits(315)	1e-51	Compositional matrix adjust.	112/229(49%)	135/229(58%)	7/229(3%)	-2

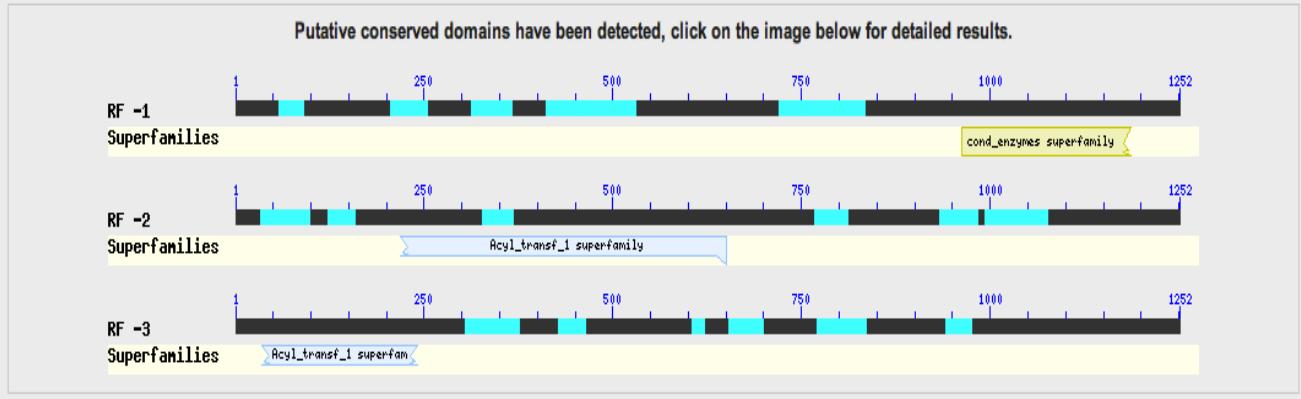
```
Query  884   AVTPAALSEQAARLVDFLGATGDValaeiagallrgraaLDHRAVVLAGAVDDLIRGLTD  705
             A T  AL  QA RL+D++      +  A +A ALL  RA  DHRA V+       L+ GL
Sbjct  485   AHTEPALRAQAERLLDYVTEHPGLTPAAVACALLGSRATFDHRAAVIGARPQGLLNGLDA  544

Query  704   VAAGAPG---LTGSRVNG--RVAFVFPGNRVPSGVGMGRELLVSDSVFATRLAECAAVMD  540
             ++ GAP    +TG  +    R  FVFPG       VGM  ELL +  VFA  + ECA  +
Sbjct  545   LSRGAPSADVVTGVALEAGDRPVFVFPGQGS-QWVGMATELLDTSPVFADTVRECADALA  603

Query  539   PLTGWSLLDFVAGPQEL-LDRVDVVQPVSFAVMVSLAAVWESCGVTPDVVVGHSQGEIaa  363
             P T +SLLD +        LDRVDVVQP  FA+MV LA +W S GV P  VVGHSQGEIAA
Sbjct  604   PYTNFSLLDVLRDRDPAALDRVDVVQPALFAMMVGLAELWRSHGVEPAAVVGHSQGEIAA  663

Query  362   acvagalsladaAKVVVARSRVIAQRLAGHGGMVSIALPSERVDSADRP  216
             A VAGALSL+DAA VV ARSRVI       GGM S+ALP++ V     +P
Sbjct  664   AYVAGALSLSDAALVVAARSRVIGASARRGGGMASVALPADTVREKIQP  712
```

Figure 19: Amino acid sequence similarity alignment (BLAST) of RMP 2.31. (A) The partial amino acid sequence of the gene interrupted in the non-producing mutant strain RMP 2.31 was compared with part of the amino acid sequence of the polyketide synthase from *Streptomyces roseosporus*. (B) The amino acid sequence of the gene interrupted in the non-producing mutant RMP 2.31 aligned with a partial amino acid sequence of the polyketide synthase from *Rhodococcus jostii*

There are many PKS similarity hits for RMP 2.31 (Appendix). When blasting the sequence, conserved domains within PKS regions are also identified by NCBI (Figure 20A). Acyltransferases (AT) are one of the enzymatic domains that is involved in producing polyketides. An AT is one of the three minimal modules required for a type I polyketide

synthase. The acyltransferase domain is responsible for the selection of an appropriate building block in each chain elongation cycle of the PKS (Dunn et al. 2013). The amino acid alignment of RMP 2.31 to the AT conserved domain has significant similarity (Figure 20B).

A



B

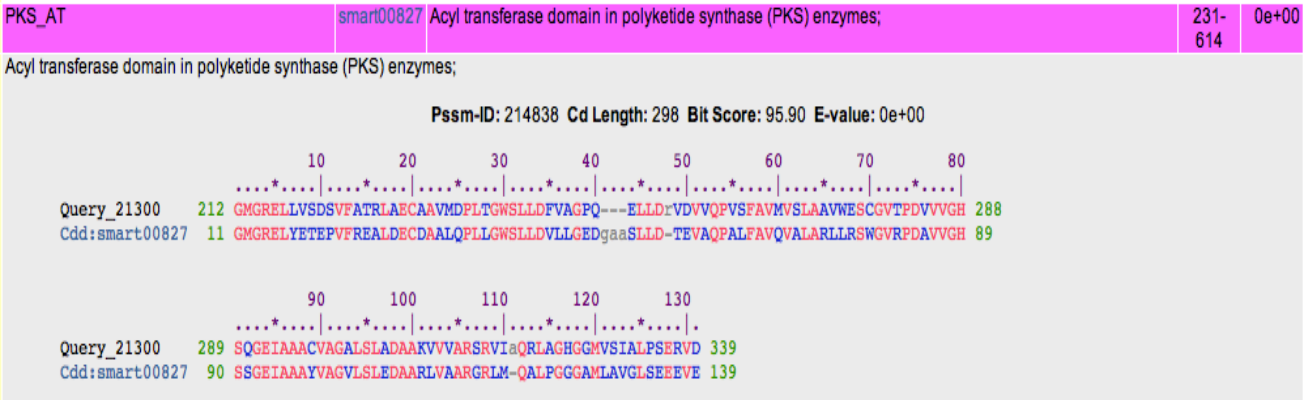
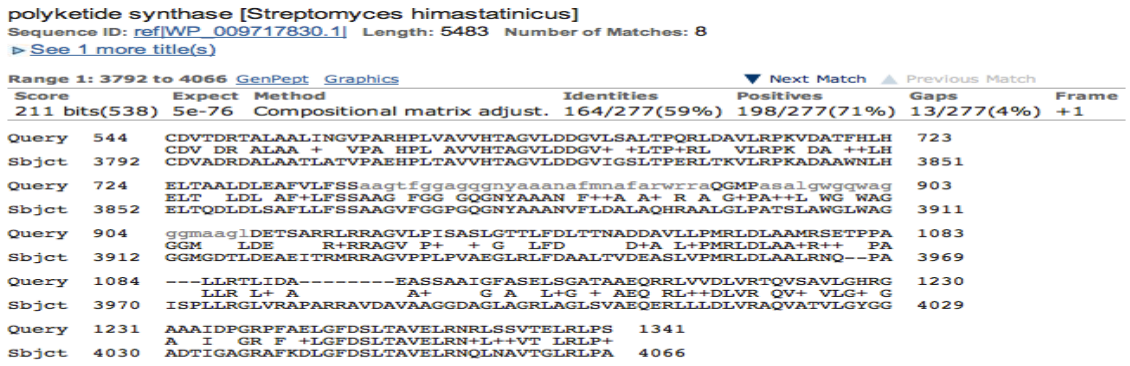


Figure 19: Amino acid conserved sequence similarity alignment (BLAST) of RMP 2.31. (A) Putative conserved domains of known polyketide synthases with the amino acids of the non-producing mutant strain RMP 2.31. (B) The *Rhodococcus* query sequence is 71% similar to the conserved functional domain acyl transferase (AT), a common module of large multi-domain polyketide synthases

RMP 77.23 is the second non-producing mutant that was identified as having a mutation in a PKS gene. It has a greater than 70% similarity to known polyketide synthase

regions within *Streptomyces* (Figure 21A) as well. The amino acid alignment for RMP 77.23 is different from that of RMP 2.31, and they show similarities to different PKS regions in different bacteria, including *Rhodococcus* (Figure 21B). As with RMP 2.31, the sequencing data align with a large number of polyketide synthase regions among bacteria (Appendix).

A



B

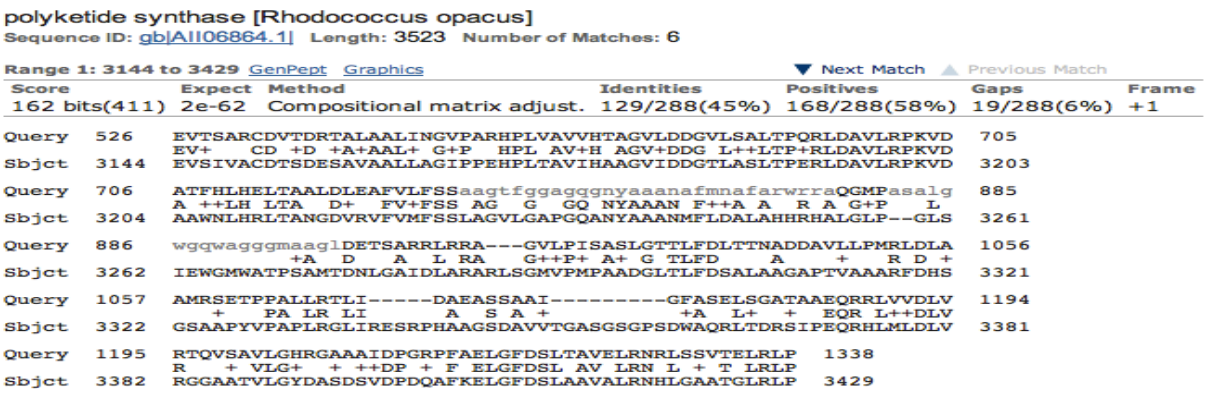
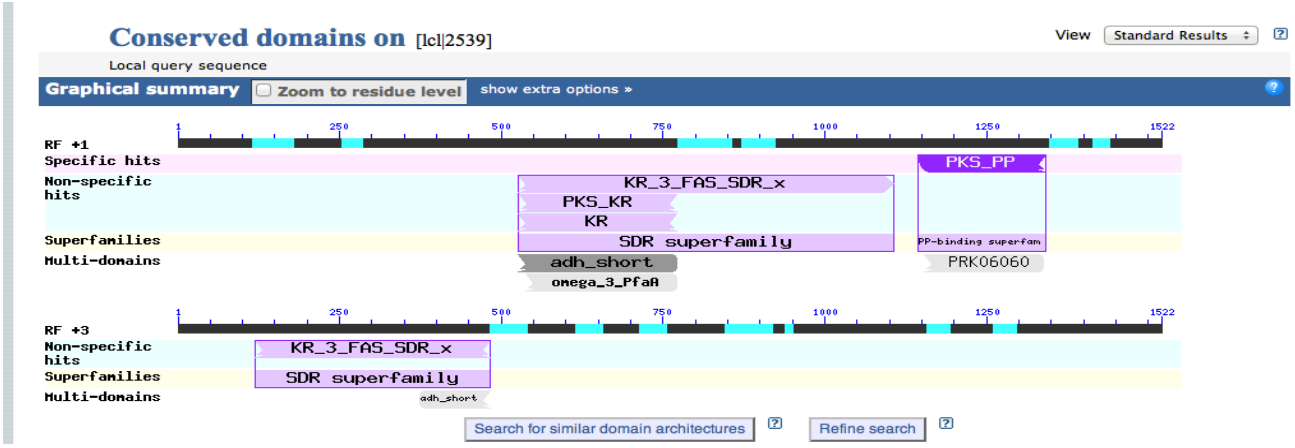


Figure 20: Amino acid sequence similarity alignment (BLAST) of 77.23. (A) The partial amino acid sequence of the gene interrupted in the non-producing mutant strain RMP 77.23 was compared with part of the amino acid sequence of a polyketide synthase from *Streptomyces himastatinicus*. (B) The amino acid sequence of the gene interrupted in the non-producing mutant strain RMP 77.23 was compared to part of the amino acid sequence of a polyketide synthase of *Rhodococcus opacus*

Like RMP 2.31, when the amino acid sequence of RMP 77.23 is blasted in the NCBI database, it shows conserved domains within known PKS regions (Figure 22A). For this mutant though, it has conserved domains in the enzymatic ketoreductase (KR) portion of the PKS that is responsible for modifying the β-carbonyl. The ketoreductase uses NADPH to reduce the keto group to a hydroxyl group (Zheng et al. 2010). The amino acid sequence also aligned with the conserved portion of phosphopantetheine, which is a prosthetic group of the acyl-carrier protein (ACP). Phosphopantetheine aids the PKS by acting as a swinging arm for attachment of amino acids. The length of the phosphopantetheine also allows intermediates to reach distant enzyme active sites (Johnson et al. 2014). The amino acid alignment of RMP 77.23 was significantly similar to that of the conserved domains (Figure 22B).

A



B

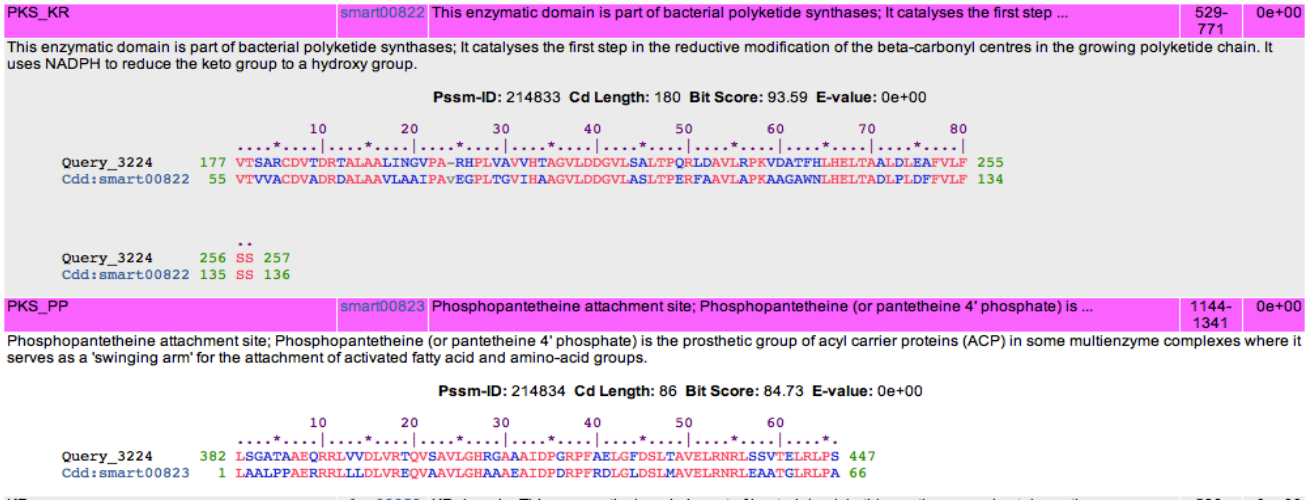


Figure 21: Amino acid conserved sequence similarity alignment (BLAST) of 77.23. (A) Putative conserved domains of known polyketide synthases with the amino acids of the non-producing mutant strain RMP 77.23. (B)The *Rhodococcus* sequence is 71% similar to the conserved functional domain beta-ketoacyl reductase (KR), a common domain of large multi-module polyketide synthases

The other 6 non-producing mutants were sequenced as well. RMP 71.7 and RMP 71.3 have an interruption in the cell division FtsK/SpoIIIE gene (Appendix). The similarity between the amino acid alignment of RMP 71.1 is 70%, and this alignment is at a different location in the FtsK/SpoIIIE gene than that of RMP 71.3. The disk assay showed that the

mutation was still producing the inhibitor similar to RMP 71.3, but at a lower concentration than that of the producer strain (Appendix).

There was no zone of inhibition observed by the mutation of RMP 46.37 in the disk diffusion assay. When the interrupted gene was sequenced, the amino acids aligned 100% similarity to a putative GntR family transcriptional regulator in *Rhodococcus erythropolis* (Appendix).

RMP 46.43, like two of the previously described non-producers, has an interruption in the FtsK/SpoIIIE gene(s) involved in cell division. There is a 77% similarity between the amino acids of the RMP 46.43 mutation and the Fts/SpoIIIE gene (Appendix). Like the previous FtsK/SpoIIIE mutants, there is a small zone of inhibition on the disk diffusion assay (Appendix). Even though RMP 71.3, RMP 71.7 and RMP 46.43 aligned with the same cellular function, these mutations were in different genes as suggested by the different locations of hybridization on the Southern blot.

RMP 82.5 aligned 57% with a beta-glucosidase gene(s) found in *Streptomyces fradie* (Appendix).

HPLC Analysis of Producer and Non-producer

Sephadex LH-20 Column Chromatography

As previously described, each scale-up was achieved by extracting 15 large RM agar plates with MTM3W5.2, RMP 2.31 or RMP 77.23 after two weeks of incubation at 19°C. The dried down extract was re-dissolved in 6 mL of methanol and centrifuged for 30 minutes to remove any solid particulates. The extract at this point was a yellow liquid. The extract was

then passed through a LH-20 column, and after collection of fractions from the column; the crude extract was a clear color. The active fractions (fractions 13 – 31) were pooled together and dried down. The dried residue was then re-dissolved in 7 mL of 90% methanol, and was then further purified using high-pressure liquid chromatography (HPLC). The mutant extracts were run through the LH-20 column exactly like the producer strain. Fractions 17-34 were collected and tested for activity using the disk diffusion assay. No activity was observed for RMP 2.31 or RMP 77.23. Fractions 13 – 31 were pooled together and dried down like the parental strain.

The re-dissolved extract was filtered using a syringe to remove any precipitate, and was then run through the HPLC program that was setup by Borisova in 2011. The program was setup so that the mobile phase began at 90% methanol, and this was gradually increased to 100% over 30 minutes. 2 mL was loaded in the HPLC column, and 35 fractions were collected and tested for inhibitory activity using the disk diffusion assay. Active fractions were present in fractions 15 – 17 (Figure 23A). The two mutant extracts were run through the HPLC using the same program as the producer strain, and fractions 15 – 17 were collected and tested for activity. There was no activity seen for either mutant (Figure 23 B & C). Fractions 1 – 30 were tested for each mutant to confirm that there wasn't any inhibitory activity from any of the fractions, and there was none found.

HPLC was done for the producer strain and the two non-producing mutants to visually show the inhibitory compound present in the producer strain, MTM3W5.2, but not seen in either of the mutants. A small peak is observed in the chromatogram for RMP 2.31 (Figure 23B) that is in the same location as the peak for the inhibitory compound in the producer strain. The fractions for this peak were tested, and there was not any activity.

Since this is a crude extract, this chromatogram may not separate or resolve single molecules. That is, the broad peak in (A) is not a single molecule but a mix of several different molecules (only one is lost in the mutant chromatograph of RMP 2.31). Another possibility could be that the column being used for HPLC was not flushed well enough before running the mutant, and some of the producer strain's extract was left in the column. This could have been at such a low concentration that it was barely detectable and had no activity.

A

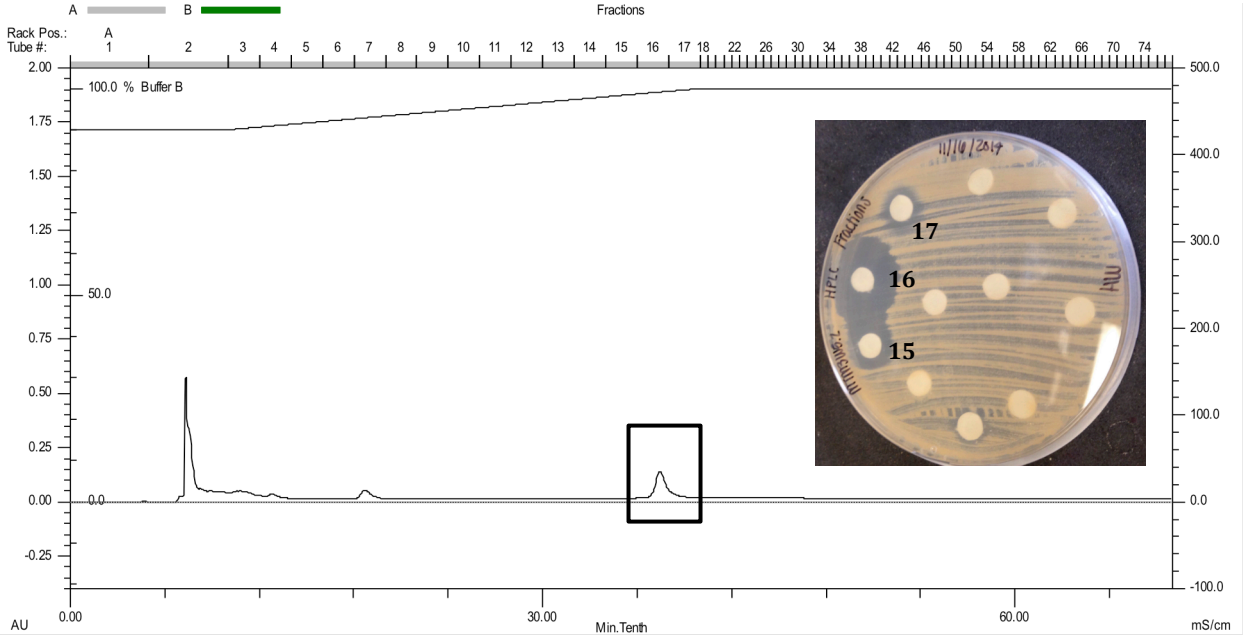
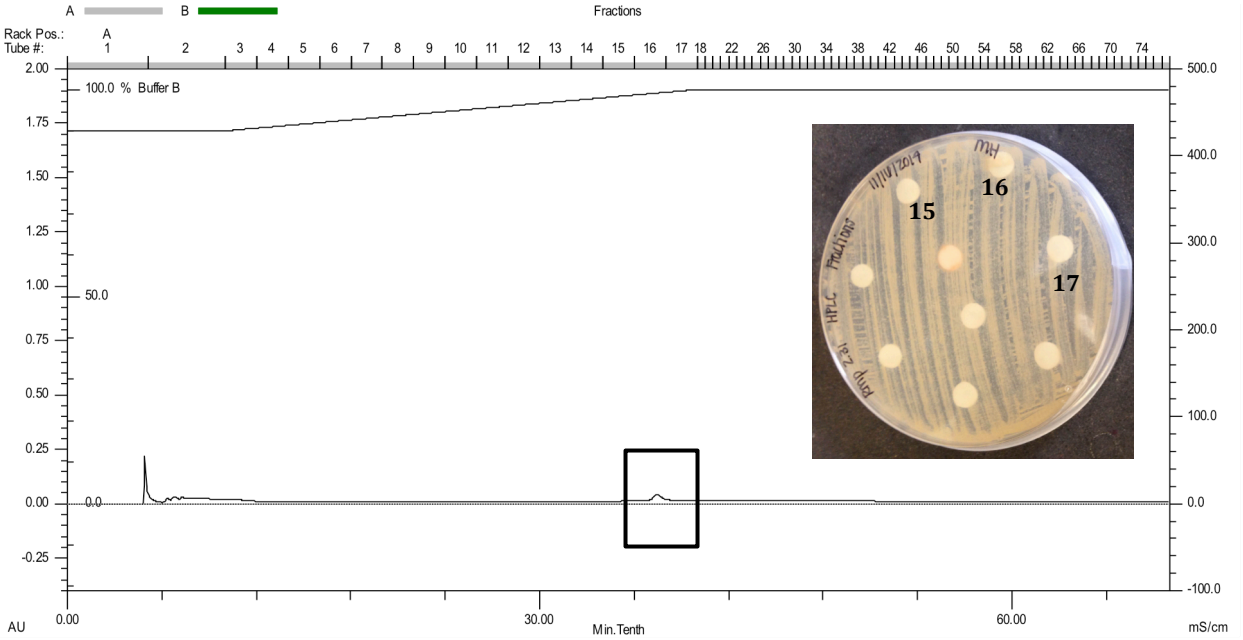


Figure 22 (continued on next page)

B



C

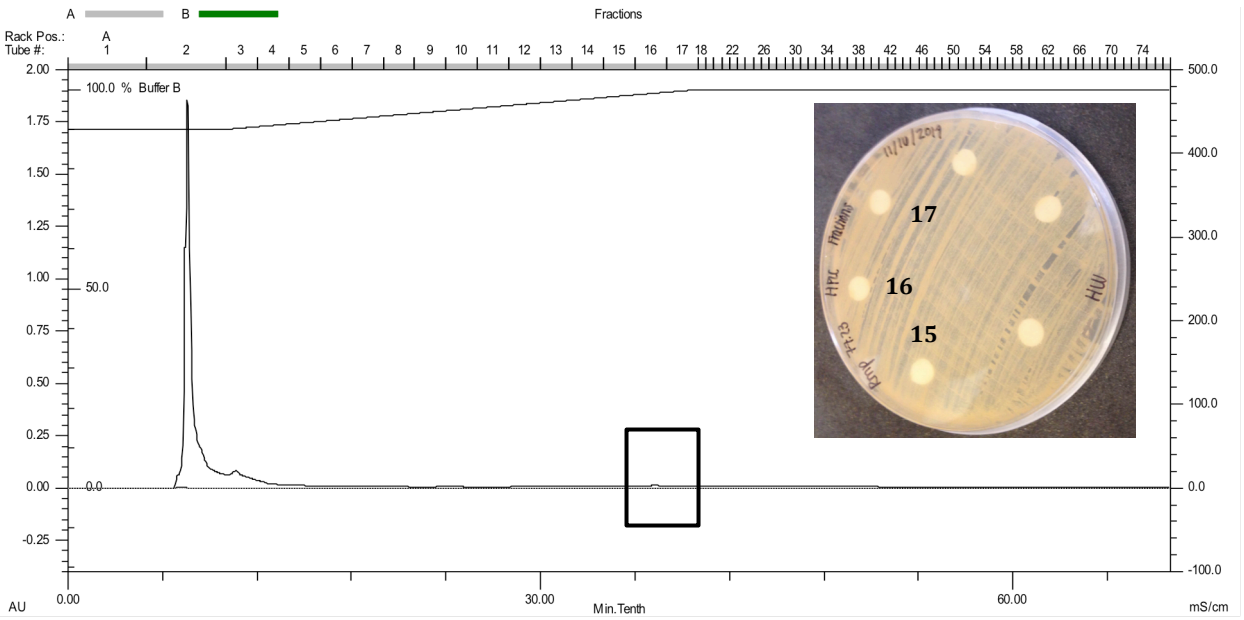


Figure 22: HPLC analysis of the inhibitory compound. HPLC analysis is shown of the agar plate culture extract from the producer strain MTM3W5.2 (A), the non-producing mutant RMP 2.31 (B), and the non-producing mutant RMP 77.23 (C). Note that inhibitory activity is only detected in column fractions 15, 16, and 17 from the parent strain MTM3W5.2; picture inset in (A) of disk diffusion assay, but no activity is observed in these same fractions from the two non-producing mutants (B & C)

CHAPTER 4

DISCUSSION

The urgency to identify novel antibiotics has led researchers to begin examining organisms that were never looked at for antibacterial, antifungal and anticancer properties. *Rhodococcus,* although always recognized for its diverse enzymatic potential, has largely been overlooked in the field of drug development. *Rhodococcus* is a genus of bacteria that is related to *Streptomyces*. *Streptomyces* is the largest producer of antibiotics used in the clinic today (Zhu 2011). Due to their close relation, investigating Rhodococci for inhibitor production is warranted, and in 2006 when the first *Rhodococcus* genome was sequenced, it was discovered that the genome had multiple PKS and NRPS genes (McLeod et al. 2006). Since 2006, there have been many other rhodococci genomes sequenced, and all of them continue to have a large number of PKS and NRPS gene clusters (Table 2). The knowledge of these enzymes that are known to produce inhibitory compounds has led to the discovery of several compounds such as a cyclic peptide structure (Iwatsuki et al. 2007), an aurachin structure (Kitagawa 2008), a rhodopeptin (Chiba et al. 1999) and a rhodostreptomycin (Kurosawa et al. 2007) among *Rhodococcus* species.

Agar Extraction Assay

Based on this knowledge, a previous graduate student screened many newly isolated *Rhodococcus* species and tested for inhibitory activity against a small library of Gram-positive and Gram-negative organisms. MTM3W5.2 was identified as producing a strong inhibitor against *Rhodococcus* and other closely related Gram-positive bacteria

(Figure 15A). By identifying the gene(s) involved in the biosynthesis of the inhibitory compound produced by *Rhodoccocus* sp. MTM3W5.2, sophisticated bioinformatic analysis tools can provide information about the chemical composition of the inhibitor.

I began the project by screening a library of MTM3W5.2 pTNR transposon mutants. These mutants were screened by the agar extraction method. Borisova (2011) had identified that this was the most efficient and reliable method to use in the screening process of this strain. The agar extraction screen was laborious but provided insight into the gene(s) responsible for the biosynthesis of the inhibitor. The inhibitor produced a large zone of inhibition (~33 mm), which made the disk diffusion assay an easy method to quickly determine if the mutant was a producer or not (Figure 15B & C). I never received a false result for a mutant that was not producing a zone of inhibition, and most of the producing mutants maintained the same zone size as the producer MTM3W5.2 strain.

Eight mutants were identified as non-producing mutants out of 2,306 Kan resistant colonies screened (Table 4). Typically in a knockout screen, it is important to saturate the genome with mutations. This means that the transposon will interrupt every gene (average genome size of *E. coli* is approximately 1.5 – 2 Kb) in the genome of interest. For the purposes of this project, we were not as concerned about saturation of the genome. The genome size of MTM3W5.2 is still unknown, and *Rhodococcus* bacteria are known to produce some of the largest genomes within bacteria (as large as 9.7 Mb in size) (McLeod et al. 2006), so the process of getting genome saturation could have taken significantly more than 2,306 mutants. If we wanted to determine the size of the MTM3W5.2 genome, we could proceed with sequencing the genome using next generation sequencing. Because the genes required to synthesize antibiotic molecules can be quite large (PKS genes

especially > 100,000 Kb) this gives the transposon a much larger "target". Thus ≥ 2,000 mutants screened can be enough to hit a gene involved in the biosynthesis of the inhibitory compound.

Auxotrophic Mutant Screen

An auxotrophic mutant screen was done with the pTNR mutant library. This was an assay that confirmed that there were actual mutations that were being created by the transposon system. The likelihood that a mutation interrupted a gene required to synthesize one of the twenty amino acids is common because there are many genes necessary for the synthesis of amino acids, and thus the reason that each mutant was tested on a minimal medium. I initially had to identify a minimal medium that the producer strain would grow on. I began with fiss minimal medium, but the producer would not grow well on this medium. It took almost two weeks to see any growth on the medium. I then tried M3 minimal medium. Although the producer grew at a much slower rate on the M3 minimal medium than compared to a rich medium, it would grow on the minimal agar plate within four to five days. These plates were grown on the bench-top. 7 mutants were identified out of the 2,306 mutants to be auxotrophic mutants (Figure 14 and Table 4). This may seem like a low number in comparison to the number of mutants screened, but Pratt (2008) detected 5 auxotrophs out of 2,331 pTNR transformants and Vellore (2001) detected 5 from a screen of 250 pJCS505 transformants.

Southern Blot Analysis of 8 Non-producing Mutants

Once the 8 non-producing mutants had been confirmed as non-producers, I performed a Southern blot. There were difficulties with getting enough genomic DNA extracted from each of the mutants in order to get enough transferred onto the membrane for hybridization. One issue that I ran into was that a few of the mutants grew at a significantly slower rate than the parental strain. A growth curve could be done on each of the mutants to obtain information on its growth behavior, but this has yet to be done. The standard protocol for genomic isolation was altered to grow the bacterial seed for two days instead of one, and once the cultures were inoculated with the seed, they grew for two to three days depending on the growth rate. Some of the mutants required up to four days of incubation because of such a slow growth rate.

Getting enough genomic DNA transferred to the membrane was a challenge at first, but just one mutant did not hybridize with the probe on the membrane (Figure 16B, lane 2). A possibility for lack of hybridization with the mutant could be that there is an issue with the pTNR probe. It could be too large and cannot properly hybridize to the insertion region, or the mutant does not have a pTNR insertion as expect. The latter seems less likely because the mutant continues to grow on kanamycin impregnated agar plates. A wild-type strain that does not have an insertion site from pTNR cannot do this.

The Southern confirmed that I identified 7 mutants that had insertions of the pTNR transposon in their chromosomes (1 mutant did not hybridize to the probe). This blot also showed that the transposon insertions were in different locations in the genome because the hybridized bands migrated at different positions in the agarose gel after electrophoresis. If the fragments had been at or around the same molecular weight, it

would have been possible that the pTNR transposon had inserted into the same gene within the MTM3W5.2 genome. Also, the hybridization in the Southern does not support the information given from the plasmid gel (Figure 18). The plasmid for each mutant is too small on the agarose gel in comparison to the signals that are seen on the southern. This is known because the size of the pTNR insertion site is approximately 2.7 kb, and many of the mutants on the plasmid gel are approximately 2.3 kb. A new plasmid prep needs to be obtained and a new gel run to see if there was an error with the first gel.

Sequencing Non-Producing Mutants

To identify the gene(s) that had been interrupted by the pTNR transposon, the insertion sites were sequenced using primers that had been designed to sequence from the insertion termini of the pTNR transposon into the gene of interest (Appendix). Of the 8 mutants, one mutant, RMP 70.47, does not have any sequencing data. Multiple attempts to sequence this mutant were made and all were unsuccessful. The size of the RMP 70.47 cloned plasmid was close to that of several of the other mutants (Figure 18, lane 1). A possibility that there was no sequencing data could be that the samples that were sent to be sequenced were contaminated with a reagent such as ethyl alcohol, and this could have disrupted the sequencing as well. The remaining 7 mutants produced enough DNA sequence information to predict the gene that is interrupted in these insertion mutations (Table 6) when "blasted" in the NCBI database. Only one clone was sequenced for each mutant. For future work, at least 10 clones of each mutant need to be sequenced to confirm this data.

DNA and amino acid sequences were aligned using BLAST against the NCBI database. DNA sequences from five non-producing mutants did not show similarity to known polyketide synthase regions. However, they did align with genes that are responsible for other cellular functions. RMP 71.3 aligned 62% with the FtsK/SpoIIIE gene(s) involved in cell division (Appendix). An interruption in a gene that is involved in cell division may have interrupted a function that wouldn't allow the inhibitory to be released from the cell or even be synthesized because the cell isn't functioning properly. By observing the disk assay plate, the mutant was identified as one that produces a significantly smaller zone of inhibition compared to that of the producer, MTM3W5.2 (Appendix). This suggests that the inhibitor is still being produced in the cell, but can't sufficiently be excreted, from the cell, or the inhibitor is being synthesized at a significantly lower rate than normal. This mutant is one that is a slow grower when growing it under optimal conditions. The delay in growth would explain the poor production of the compound, and thus smaller zone on the disk diffusion assay.

Like RMP 71.3, RMP 71.7 also showed similarity to the FtsK/SpoIIIE gene(s). This mutant aligned to a different sequence within the gene(s) compared to that of RMP 71.3. The sequence similarity was 70% (Appendix). The disk diffusion assay showed that the mutant produced the inhibitory compound, but at a much lower concentration than the producer strain (Appendix). The reasoning for these results is the same as RMP 71.3.

Another FtsK/SpoIIIE interruption identified 77% with the amino acid alignment of RMP 46.43 (Appendix). The disk diffusion assay for this mutation however, showed that there is a small zone of inhibition, but it is a smaller zone than that of RMP 71.3 and RMP 71.7 (Appendix). The gene(s) or partial gene(s) that were interrupted in this mutation

could have inhibited the biosynthesis or efflux of the inhibitor more than the RMP 71.3 and RMP 71.7 mutations, or it was due to the slowed growth of the mutant. RMP 46.43 aligned with a different amino acid sequence than that of RMP 71.3 and RMP 71.7.

RMP 46.37's amino acid alignment showed 100% similarity to the GntR family of transcriptional regulators (Appendix). This family of transcriptional regulators is responsible for regulating a wide variety of functions among bacteria including the production of antibiotics and on ABC transporter in *Streptomyces* (Hillerich and Westpheling 2006). The disk diffusion assay suggests that the interruption in this gene stops the synthesis of the inhibitor as observed by no zone of inhibition against the indicator (Appendix).

RMP 82.5 aligned 43% with β-glucosidase. β-glucodidases are found widely among living organsims, and are known to catalyze the hydrolysis of β1-4 bonds linking two glucose or two substituted glucose molecules (Mattéotti et al. 2011). There was no inhibitor detected in the disk diffusion assay suggesting that the biosynthesis was completely blocked by the mutation. A test could be done on this mutant to lyse the cell and test the inside of the cell to see if the inhibitor is being produced or not. This would provide insight into whether the biosynthesis is being blocked and no compound is being produced, or that there is something keeping the compound from being exported from the cell. The number of sequenced base pairs obtained for RMP 82.5 was quite low (423 bp of sequenced DNA), thus it could be possible that the interrupted gene is not even a β-glucosidase, and that there needs to be more sequencing data to get a better idea of the gene(s) that was interrupted.

It is important to note that there is small sequencing data available for each of these mutants, especially the six mutants that did not show similarity to PKS genes. To better understand and identify the gene(s) that have been interrupted in each of these mutants, more sequencing data would need to be obtained.

RMP 2.31 and RMP 77.23 were identified as having significant similarity to known polyketide synthase genes in bacteria such as *Streptomyces* as well as known PKS genes within *Rhodococcus*. Because these two mutations are located in PKS regions, they were chosen to investigate further. RMP 2.31's amino acid alignment showed conserved regions in the acyl transferase domain found in PKS. The acyl transferase is one of the basic domains that is found in polyketide synthases. The acyl transferase domain of the modular PKS has a tight control over the choice of the α-carboacyl-CoA substrate. Although the mechanism that AT undergoes within the PKS is unknown. These results are encouraging because that confirm that the mutation in in a PKS gene, and the large number of hits that are received when "blasting" the amino acid sequences provides encouragement to support this. When initially "blasting" the DNA sequence, the alignment shows significant homology to many *Streptomyces* strains (Appendix), but if you limit the search to only *Rhodococcus* strains, there is also a large number of hits to PKS regions found in rhodococci. It is important to note that the alignment to *Streptomyces* is more similar (>70%) than that of the *Rhodococcus* species ($\leq$ 50%). This may suggest a couple of things. First, since the similarity to *Rhodococcus* species is less than 50%, it could mean that this is a novel compound produced within *Rhodococcus*, and that it is possible that the compound that MTM3W5.2 is producing is more closely related to that of the *Streptomyces* strains, and

could possibly be an antibiotic. Another explanation could be that this strain has obtain some genes from *Streptomyces* through horizontal gene transfer.

RMP 77.23 is the second non-producing mutant that showed similarity to known PKS genes, as well. This mutant has conserved domains of the β-ketoacyl reductase. KRs are responsible for adding to the β-carbonyl. This sequence also had a conserved domain in one of the prosthetic groups found on the ACP. Like RMP 2.31, the sequence showed significant similarity to known *Streptomyces* PKS regions as well as those in *Rhodococcus.* The results are similar in that the similarity to *Streptomyces* is greater than 70% and to *Rhodococcus* is less than 50%.

Both of these mutants did not produce a zone on inhibition in the disk diffusion assay. The plasmids made from the cloned insertion fragment show that their plasmid size is around 4,400 bp for RMP 77.23 and 4,290 bp for RMP 2.31. The insertion size is roughly 2,600 bp, and there has only been a little over 1,000 bases sequenced for each mutant. This means that there is still a significant amount to continue sequencing to get the entire sequence of the interrupted gene(s).

Table 6: Identification of the interrupted gene in each non-producing mutant

Non-producing Mutant	**Sequenced DNA (No. of base pairs)[a]**	**Similarity to known gene[b]**
RMP 2.31	1248	Polyketide Synthase – AT domain
RMP 71.3	832	Cell Division (Ftsk/SpoIIIE)
RMP 71.7	368	Cell Division (Ftsk/SpoIIIE)
RMP 46.37	204	Putative GntR Transcriptional Regulator
RMP 46.43	371	Cell Division (Ftsk/SpoIIIE)
RMP 70.47	No sequencing data	-----
RMP 77.23	1549	Polyketide Synthase – KR domain
RMP 82.5	423	Beta-glucosidase

[a]Total number of chromosome DNA bases sequenced on both flanks of the inserted transposon.
[b] Based on BLASTx sequence alignments

HPLC Analysis of MTM3W5.2 and Two Non-producers

I did HPLC analysis on the producer, MTM3W5.2, and the two non-producers that showed interruptions in PKS genes. These two mutants were RMP 2.31 and RMP 77.23. Borisova (2011) optimized the HPLC conditions for the producer strain, and this is the program that I used. Before running the samples through HPLC, the producer and the two non-producers were scaled-up by growing them on 15 large agar plates. This protocol was also used from Borisova. The extracts were run through an LH-20 column prior to HPLC.

The purpose of doing HPLC was to see the active peak (molecule) in the producer strain, but when the two non-producing mutants were run through the HPLC, the inhibitory peak (molecule) would go away because these mutants do not produce the compound. In RMP 2.31, there was a small peak that was identified where the active peak is located in the producer. These fractions were collected and tested, but there was not any activity found. An even smaller peak was identified in the same location on the chromatogram for RMP 77.23. Again, the fractions were collected and tested by the disk diffusion assay, but no activity was observed.

The peaks could be some compound that is present in the bacteria that comes off of the column at the same time that the inhibitor does, but does not have activity. This could be possible because the HPLC program was not optimized to get the best separation of the compounds present in the extract, only enough to clearly identify the active fractions. Another possibility is that the producer, MTM3W5.2, was the first extract to be run through the HPLC system. If the column was not flushed enough after use, some of the compound could have been left behind in the column, but after flushing the column it lost its activity. It still came off of the column at the appropriate time, but it was not longer active. This would support why there is more of in seen in the RMP 2.31 chromatogram verses RMP 77.23 because RMP 2.31 was the next extract run after the producer.

Future Work

Two mutants have been identified as having interruptions in a PKS gene(s). For future work with *Rhodococcus* sp. MTM3W5.2, the entire sequence of the PKS gene cluster needs to be identified. Once the sequence is known, bioinformatic analysis tools such as

antiSMASH can be used to identify unique gene clusters and possibly predict the structure of the inhibitory compound. By knowing the genes that make up the PKS gene cluster, it could also be possible to piece together the modules and identify the number and order of modules that are responsible for the biosynthesis of the compound.

PKS gene clusters can be rather large in bacteria. After these genes were identified, the next steps were to continue sequencing out into the gene to get the entire sequence of the PKS by primer walking. Primers were designed based on the latest sequencing data to continue sequencing out into the interrupted gene in order to obtain more sequence of the gene of interest. Due to time and cost, sequencing the two mutants was stopped after obtaining a little over 1,000 bases per mutant in one direction (RMP 2.31 forward, RMP 77.23 reverse) (Appendix). It was decided that the best, and most likely cheaper way to identify the PKS sequence is to sequence the entire *Rhodococcus* MTM3W5.2 genome. By sequencing the genome using next generation sequencing methods, the PKS region can be easily identified. Also, since this is most likely a novel *Rhodococcus* strain, due to there being so few rhodococci species that have been sequenced to date, sequencing the entire genome can be beneficial for more reasons just than to obtain the PKS region. Obtaining the entire genome sequence of this strain would allow us to identify other enzymes that may be beneficial in other investigations. Also, as stated, few *Rhodococcus* genomes have been sequenced, and a new sequence would add to the current data and can be shared by researchers all over the world.

There are also 6 mutants that have decreased inhibitory compound production due to other genes that are not PKS. These mutants can also be sequenced to identify the genes that are interrupted. This data would provide insight into functions in the cell that inhibit

the production or release of an inhibitory compound aside from the biosynthetic gene cluster.

Also, this compound has a very narrow spectrum of activity. Thus far, it is known that the compound inhibits *Rhodococcus* and other closely related Gram-positive genera. The DNA and amino acid alignments of the genes responsible for synthesizing the inhibitor align with well-known antibiotics. It is possible that the compound could be an antibiotic that inhibits a specific unknown bacterium. It would be beneficial to screen this inhibitor against a larger library of bacteria to confirm its spectrum of activity.

Along with this work, a collaborative effort has been initiated to begin examining the inhibitory compound through HPLC, mass spec and proton NMR. Through these efforts, it is possible that the compounds characteristics can be determined such as the molecular formula, structure and size.

Once the genes and structure of the compound is known, the project could be taken further to examine the mechanism of action of the inhibitor. A problem with antibiotics that are used clinically is that they are all targeting the same areas of a cell. Identifying the target of this inhibitor would be important in the advancement of this potentially novel compound.

REFERENCES

Alvarez HM, Silva RA, Cesari AC, Zamit AL, Peressutti SR, Reichelt R, Keller U, Malkus U, Rasch C, Maskow T, Mayer F, Steinbuchel A. 2004. Physiological and morphological responses of the soil bacterium *Rhodococcus opacus* strain PD630 to water stress. FEMS Microbiol Ecol. 50(2): 75–86.

Auffret M, Labbe D, Thouand G, Greer CW, Fayolle-Guichard F. 2009. Degradation of a Mixture of Hydrocarbons, Gasoline, and Diesel Oil Additives by *Rhodococcus aetherivorans* and *Rhodococcus wratislaviensis*. Appl Environ Microbiol. 75(24): 7774–82.

Bala M, Kumar S, Raghava GPS, Mayilraj S. 2013a. Draft Genome Sequence of *Rhodococcus qingshengii* Strain BKS 20-40. Genome Announc*.* 1(2): 128–33.

Bala M, Kumar S, Raghava GPS, Mayilraj S. 2013b. Draft Genome Sequence of *Rhodococcus ruber* Strain BKS 20-38. Genome Announc*.* 1(2): 139–53.

Bell KS, Philip JC, Aw DWJ, Christofi N. 1998. The genus *Rhodococcus.* J Appl Microbiol. 85(2): 195–210.

Bérdy J. 2012. Thoughts and facts about antibiotics: Where we are now and where we are heading. J Antibiot. 65(8): 385-95.

Boddy CN. 2013. Bioinformatics tools for genome mining of polykeptide and non-ribosomal peptides. J Indl Microbiol Biotechnol*.* 41(2): 443–50.

Borisova R. 2011. Isolation of a *Rhodococcus* Soil Bacterium that Produces a Strong Antibacterial Compound [Thesis]. [Johnson City (TN)]. East Tennessee State University.

Brooks SL, Van Hamm JD. 2012. Whole-Genome Shotgun Sequence of *Rhodococcus* Species Strain JVH1. J Bacteriol. 194(19): 5492-3.

Cane DE, Walsh CT. 1999. The parallel and convergent universes of polyketide synthases and nonribosomal peptide synthetases. Chem Biol*.* 6(12): 319–25.

Carr M. 2012. Detection and Purification of a Novel Natural Inhibitory Compound from an Isolated Strain of *Rhodococcus* Using an Agar Extraction Method [Thesis]. [Johnson City (TN)]: East Tennessee State University.

Center for Disease Control. 2013. Antibiotic Resistance Threats in the United States. Atlanta (GA): Center for Disease Control; [accessed 2015 Jan 25]. http://www.cdc.gov/drugresistance/threat-report-2013/pdf/ar-threats-2013-508.pdf.

Cheng YQ, Tang GL, Shen B. 2003. Type I polyketide synthase requiring a discrete acyltransferase for polyketide biosynthesis. Proc Natl Acad Sci USA. 100(6): 3149–54.

Chiang YM, Chang SL, Oakley BR, Wang CC. 2011. Recent advances in awakening silent biosynthetic gene clusters and linking orphan clusters to natural products in microorganisms. Curr Opin Chem Biol. 15(1): 137–43.

Chiba H, Agernatu H, Kaneto R, Terasawa T, Sakai K, Kazuyuki D, Yoshioka T. 1999. Rhodopeptins (Mer-N1033), novel cyclic tetrapeptides with antifungal activity from *Rhodococcus* sp. 1. taxonomy, fermentation, isolation, physio-chemical properties and biological activities. J Antibiot (Tokyo). 52(8) 695–9.

Coates ARM, Halls G, Hu Y. 2011. Novel class of antibiotics or more of the same?. Br J Pharmacol. 163(1): 184–94.

Davies J, Davies D. 2010. Origins and evolution of antibiotic resistance. Microbiol Mol Biol Rev. 74(3): 417–33.

Dias DA, Urban S, Roessner U. 2012. A historical overview of natural products in drug discovery. Metabolites 2(2): 304–36.

Dueholm MS, Albersten M, D'Imperio S, Tale VP, Lewis D, Nielsen PH, Nielsen JN. 2014. Complete Genome of *Rhodococcus pyridinivorans* SB3094, a Methyl-Ethyl-Ketone-Degrading Bacterium Used for Bioaugmentation. Genome Announc. 2(3): 525–32.

Dunn BJ, Cane DE, Khosla C. 2013. Mechanism and Specificity of an Acyltransferase Domain from a Modular Polyketide Synthase. Biochemistry. 52(11): 1839–41.

Euzéby JP. (2015 Mar 15). List of Prokaryotic Names with Standing in Nomenclature. 2015. – Genus *Rhodococcus*. Retrieved from http://www.bacterio.cict.fr/qr/rhodococcus.html.

Fischbach MA, Walsh CT. 2009. Antibiotics for emerging pathogens. Science. 325(5944): 1089–93.

Goodfellow M. 1989. Bergey's Manual of Systematic Bacteriology. Philadelphia, PA: Williams and Wilkins. Section 26: nocardioform actinomycetes – Genus *Rhodococcus*. Vol. 4. Baltimore (MD). Wilkins and Williams.

Gross H, Stockwell V, Henkels MD, Nowak-Thompson B, Loper JE, Gerwick WH. 2007. The Genomisotopic Approach: A Systematic Method to Isolate Products of Orphan Biosynthetic Gene Clusters. Chem Biol. 14(1): 53–63.

Haynes SW, Challis GL. 2007. Non-linear enzymatic logic in natural product modular mega-synthases and synthetases. Curr Opini Drug Discov Devel. 10(2): 203-10.

Hillerich B, Westpheling J. 2006. A New GntR Family Transcriptional Regulator in *Streptomyces coelicolor* Is Required for Morphogeneis and Antibiotic Production and Controls Transcription of an ABC Transporter in Response to Carbon Source. J Bacteriol. 188(21): 7477-87.

Hjerde E, Pierechod MM, Williamson AK, Bjerga GEK, Willassen NP, Smalâs AO, Altermark B. 2013. Draft Genome Sequence of the Actinomycete *Rhodococcus* sp. Strain AW25M509, Isolated From the Hadsel Fjord, Northern Norway. Genome Announc. 1(2): 1.

Inokoshi J, Matsuhama M, Miyake M, Ikeda H, Tomoda H. 2012. Molecular cloning of the gene cluster of lariatin biosynthesis of *Rhodococcus jostii* K01-B0171. Appl Microbiol Biotechnol. 95(2): 451-60.

Iwabuchi N, Sunairi M, Urai M, Itoh C, Anzai H, Nakajima M, Harayama S. 2002. Extracellular polysaccharides of *Rhodococcus rhodochrous* S-2 stimulate the degradation of aromatic components in crude oil by indigenous marine bacteria. Appl Environ Microbiol. 68(5): 2337–43.

Iwatsuki M, Uchida R, Takakusagi Y, Matsumoto A, Jiang C, Takahashi Y, Arai M, Kobayashi S, Matsumoto M, Inokoshi J, Tomoda H, Omura S. 2007. Lariatins, Novel antimycobacterial peptides with a lasso structure, produced by *Rhodococcus jostii* K01-B0171. J Antibiot (Tokyo). 60(6): 357–63.

Johnson MNR, Londergan CH, Charkoudian LK. 2014. Probing the Phosphopantetheine Arm Conformations of Acyl Carrier Proteins Using Vibrational Spectroscopy. J Amer Chem Soc. 136(32): 11240–3.

Kitawaga W, Tamura T. 2008a. Three types of antibiotics produced from *Rhodococcus erythropolis* strains. Microbes Environ. 23(2): 167–71.

Kitawaga W, Tamura T. 2008b. A quinolone antibiotic from *Rhodococcus erythropolis* JCM 6824. J Antibiot (Tokyo). 61(11): 680–2.

Kobayashi M, Shimizu S. 1998. Versatile nitrilases: nitrile-hydrolyzing enzymes. FEMS Microbiol Lett. 120(3): 217–24.

Kumar S, Bala M, Raghava GPS, Mayilraj S. 2013. Draft Genome Sequence of *Rhodococcus triatomae* Strain BKS 15-14. Genome Announc. 1(2): 128–33.

Kurosawa K, Ghivirgia I, Sambandan TG, Lessard PA, Barbara JE, Rha C, Sinskey AJ. 2007. Rhodostreptomycins, antibiotics biosynthesized following horizontal gene transfer from *Streptomyces padanus* to *Rhodococcus facians*. J Amer Chem Soc. 130(4): 1126–7.

LeBlanc JC, Goncalves ER, Mohn WW. 2008. Global response to desiccation stress in the soil actinomycete *Rhodococcus jostii* RHA1. Appl Environ Biol. 74(9): 2627–36.

Letek M, Gonzalez P, MacArthur I, Rodriguez H, Freeman T, Valero-Rello A, Blanco M, Buckley T, Cherevach I, Fahey R, Hapeshi A, Holdstock J, Leadon D. 2010. The Genome of a Pathogenic *Rhodococcus*: Cooperative Virulence Underpinned by Key Gene Acquisitions. PLoS Genet. 6(9): 1–17.

Liu B, Pop M. 2009. Antibiotic Resistance Gene Database. Nucleic Acids Res. 37(Database issue): D443-7.

Matteotti C, Thonart P, Francis F, Haubruge E, Destain J, Brasseur C, Bauwens J, De Pauw E, Portetelle D, Vandenbol M. 2011. New glucosidase activities identified by functional screening of a genomic DNA library from the gut microbiota of the termite *Reticulitermes santonensis*. Microbiol Res*.* 166(8): 629-42.

Medema MH, Blin K, Cimermancic P, de Jager V, Zakrzewski P, Fischbach MA, Weber T, Takano E, Breitling R. 2011. antiSMASH: rapid identification, annotation and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences. Nucleic Acids Res*.* 39(3): 339–46.

McLeod MP, Warren RL, Hsiao WW, Araki N, Myhre M, Fernandes C, Miyazawa D, Wong W, Lilquist AL, Wang D, Dosanjh M, Hara H, Petrescu A, Morin RD, Yang G, Stott JM, Schein JE, Shin H, Smailus D, Siddiqui AS, Marra MA, Jones SJ, Holt R, Brinkman FS, Miyauchi K, Fukuda M, Davies JE, Mohn WW, Eltis LD. 2006. The complete genome of *Rhodococcus* sp. RHA1 provides insights into a catabolic powerhouse. Proc Natl Acad Sci USA. 103(42): 15582–7.

Neilands JB. 1995. Siderophores: structure and function of microbial iron transport compounds. J Biol Chem. 270(45): 26723-7.

Pachebat JA, Keulan GV, Whitten M, Girdwood S, Del Sol R, Dyson PJ, Facey PD. 2013. Draft Genome Sequence of *Rhodococcus rhodnii* Strain LMG5362, a Symbiont of *Rhodnius prolixus* (Hemiptera, Reduviidae, Triatominae), the Principle Vector of *Trypanosoma cruzi*. Genome Announc*.* 1(3): 329–38.

Parsek MR, Greenberg EP. 2005. Sociomicrobiology: the connections between quorum sensing and biofilms. Trends Microbiol. 13(1): 27-33.

Partridge E, Gareiss P, Kinch MS, Hoyer D. 2015. An analysis of FDA-approved drugs: natural products and their derivatives. Drug Discov Today. 6445(15): 1–4.

Pathak A, Green SJ, Ogram A, Chauhan A. 2013. Draft Genome Sequence of *Rhodococcus opacus* Strain M213 Shows a Diverse Catabolic Potential. Genome Announc*.* 1(1): 144–53.

Penn K, Wang J, Fernando SC, Thompson JR. 2014. Secondary metabolite gene expression and interplay of bacterial functions in a tropical freshwater cyanobacterial bloom. ISME J. 8(9): 1866–78.

Pratt M. 2008. Iron Acquisition in *Rhodococcus erythropolis* Strain IGTS8: Characterization of a Mutant Strain That Over Produces Siderophore [Thesis]. [Johnson City (TN)]. East Tennessee State University.

Rajagopal L, Sundari CS, Balasubramanian D, Sonti RV. 1997. The bacterial pigment xanthomonadin offers protection against photodamage. FEBS Lett. 415(2): 125–8.

Ridley CP, Khosla C. 2007. Synthesis and biological activity of novel pyranopyrones derived from engineered aromatic polyketides. ACS Chem Biol. 2(2): 104–8.

Rowbotham TJ, Cross T. 1977. Ecology of *Rhodococcus coprophilus* and associated actinomycetes in fresh water and agricultural habitats. J Gen Microbiol. 100(2): 231–40.

Sallam KI, Mitani Y, Tamura T. 2005. Construction of random transposition mutagenesis system in *Rhodococcus erythropolis* using IS *1415*. J Bacteriol. 121(1): 13–22.

Sekine M, Tanikawa S, Omata S, Saito M, Fujisawa T, Tsukatani N, Tajima T, Sekigawa T, Kosugi H, Matsuo Y. 2006. Sequence analysis of three plasmids harboured in *Rhodococcus erythropolis* strain PR4. Environ Microbiol. 8(2): 334–6.

Seyedsayamdost MR. 2014. High-throughput platform for the discovery of elicitors of silent bacterial gene clusters. Proc Natl Acad Sci USA. 111(20): 7266-71.

Singh R, Smitha MS, Singh SP. 2014. The role of nanotechnology in combating multi-drug resistant bacteria. J Nanosci Nanotechnol. 14(7): 4745-56.

Smith JL, Skiniotis G, Sherman DH. 2015. Architecture of the polyketide synthase module surprises from electron cryo-microscopy. Curr Opin Struct Biol. 31: 9-19.

Spellberg B, Guidos R, Gilbert D, Bradley J, Boucher HW, Scheld WM, Bartlett JG, Edwards JJ. 2008. The Epidemic of Antibiotic –Resistant Infections: A Call to Action for the Medical Community from the Infectious Disease Society of America. Clin Infect Dis. 46(2): 155–64.

Tang L, Fu H, McDaniel R. 2000. Formation of functional heterologous complexes using subunits from the picromycin, erythromycin and oleandomycin polyketide synthases. Chem Biol. 7(2): 77–84.

Tindall BJ. 2014. A note on the genus name *Rhodococcus* Zopf 1891 and its homonyms. Int J Syst Evol Microbiol. 64(Pt 3): 1062–4.

Vellore J. 2001. Iron Acquisition in *Rhodococcus erythropolis:* the Isolation of Mutant(s) that Do Not Produce a Siderophore [Thesis]. [Johnson City (TN)]. East Tennessee State University.

Vikram S, Kumar S, Subramanian S, Raghava GPS. 2012. Draft Genome Sequence of the Nitrophenol-Degrading Actinomycete *Rhodococcus imtechensis* RKJ300. J Bacteriol. 194(13): 3543.

Walsh CT, Wencewikcz TA. 2014. Prospects for new antibiotics: a molecule-centered perspective. J Antibiot. 67(1) 7–22.

Withers TR, Yin Y, Yu HD. 2014. Identification of Novel Genes Associated with Alginate Production in *Pseudomonas aeruginosa* Using Mini-*himar1* Mariner Transposon-mediated Mutagenesis. J Vis Exp. 10(85): 3-7.

Wright W. 2010. Isolation and identification of the siderophore “vicibactin” produced by *Rhizobium leguminosarum* ATCC 14479 [Thesis]. [Johnson City (TN)]. East Tennessee State University.

Zheng J, Taylor CA, Piasecki SK, Keatinge-Clay AT. 2010. Structure and Functional Analysis of A-Type Ketoreductases from the Amphotericin Modular Polyketide Synthase. Structure. 18(8): 913–22.

Zhang Y, Qin F, Qiao J, Li G, Shen C, Huang T, Hu Z. 2012. Draft Genome Sequence of *Rhodococcus* sp. Strain P14, a Biodegrader of High-Molecular-Weight Polycyclic Aromatic Hydrocarbons. J Bacteriol. 194(13): 3546.

Zhu F, Qin C, Tao L, Liu X, Shi Z, Ma X, Jia J, Tan Y, Cui C, Lin J, Tan C, Jiang Y, Chen Y. 2011. Clustered patterns of species origins of nature-derived drugs and clues for bioprospecting. Proc Natl Acad Sci USA. 108(31): 12943-8.

APPENDICES

Appendix A: pTNR Transposon Sequence

ID

XX

7…192 IS border (left)

52…57 Duplicated target DNA

553…1347 aminoglycoside-3'-0-phosphotransferase (Kan^r) ORF

1381…1554 terminator of dnaK from *S. coelicolor*

1570…2678 pGEM3ZM *ori* region

2679... 2850 IS border (right)

2820...2825 Duplicated target DNA

2851…2973 terminator of ThcA (complement)

2974…3765 *IstB* (complement)

3765…5306 *IstA* (complement)

5350...5473 Pnit promoter (complement)

58...2819 Insertion region (tgtcga…ttgaca)

*The pTNR insertion sequence (58 – 2819) has been **bolded**.

SQ Sequence 5473 BP; 1059 A; 1613 C; 1703 G; 1098 T; 0 other;

TGTACAgcgt ttccccatgg ccagcagtac gtgtggcgtg ggtcagcagt gagtgcg**TGT** 60

CGACGGCCAT CGTAAGCTCC CCACTGGCGG CCATTTATGG CGCATTTTGT GGCCATGGTT 120

TCTCCCCGCG TACGGCCAGA TAGTTCCCCA CCTGCTGTGA Gtcagattcg gtcccggcaa 180

cgctgggtgg atGGATCCtc tagacgctgc cgcaagcact cagggcgcaa gggctgctaa 240

aggaagcgga acacgtagaa agccagtccg cagaaacggt gctgaccccg gatgaatgtc 300

agctactggg ctatctggac aagggaaaac gcaagcgcaa agagaaagca ggtagcttgc 360

agtgggctta catggcgata gctagactgg gcggttttat ggacagcaag cgaaccggaa 420

ttgccagctg gggcgccctc tggtaaggtt gggaagccct gcaaagtaaa ctggatggct 480

ttcttgccgc caaggatctg atggcgcagg ggatcaagat cCgatcaaga gacaggatga 540

ggatcgtttc gcATGATTGA ACAAGATGGA TTGCACGCAG GTTCTCCGGC CGCTTGGGTG 600

GAGAGGCTAT TCGGCTATGA CTGGGCACAA CAGACAATCG GCTGCTCTGA TGCCGCCGTG 660

TTCCGGCTGT CAGCGCAGGG GCGCCCGGTT CTTTTTGTCA AGACCGACCT GTCCGGTGCC 720

CTGAATGAAC TGCAGGACGA GGCAGCGCGG CTATCGTGGC TGGCCACGAC GGGCGTTCCT 780

TGCGCAGCTG TGCTCGACGT TGTCACTGAA GCGGGAAGGG ACTGGCTGCT ATTGGGCGAA 840

GTGCCGGGGC AGGATCTCCT GTCATCTCAC CTTGCTCCTG CCGAGAAAGT ATCCATCATG 900

GCTGATGCAA TGCGGCGGCT GCATACGCTT GATCCGGCTA CCTGCCCATT CGACCACCAA 960

GCGAAACATC GCATCGAGCG AGCACGTACT CGGATGGAAG CCGGTCTTGT CGATCAGGAT 1020

GATCTGGACG AAGAGCATCA GGGGCTCGCG CCAGCCGAAC TGTTCGCCAG GCTCAAGGCG 1080

CGCATGCCCG ACGGCGAGGA TCTCGTCGTG ACTCATGGCG ATGCCTGCTT GCCGAATATC 1140

ATGGTGGAAA ATGGCCGCTT TTCTGGATTC ATCGACTGTG GCCGGCTGGG TGTGGCGGAC 1200

CGCTATCAGG ACATAGCGTT GGCTACCCGT GATATTGCTG AAGAGCTTGG CGGCGAATGG 1260

GCTGACCGCT TCCTCGTGCT TTACGGTATC GCCGCTCCCG ATTCGCAGCG CATCGCCTTC 1320

TATCGCCTTC TTGACGAGTT CTTCTGAgcg ggactctggg gttcgaaatg accgACTAGA 1380

cagtcggcac aagggaaacg gcagcaccgg ccccggagga ccaccgcgtc ctccggggcc 1440

gattcacaac ggtgctgtct cattggtcta gtacttcttg acgatttggt gaagacagcg 1500

ttggctctct tcaccggggt cacctgggtc acctggagag caactaacta gtaaGACCTG 1560

CAGGCATGCA AGCTTGAGTA TTCTATAGTG TCACCTAAAT AGCTTGGCGT AATCATGGTC 1620

ATAGCTGTTT CCTGTGTGAA ATTGTTATCC GCTCACAATT CCACACAACA TACGAGCCGG 1680

AAGCATAAAG TGTAAAGCCT GGGGTGCCTA ATGAGTGAGC TAACTCACAT TAATTGCGTT 1740

GCGCTCACTG CCCGCTTTCC AGTCGGGAAA CCTGTCGTGC CAGCTGCATT AATGAATCGG 1800

CCAACGCGCG GGGAGAGGCG GTTTGCGTAT TGGGCGCTCT TCCGCTTCCT CGCTCACTGA 1860

CTCGCTGCGC TCGGTCGTTC GGCTGCGGCG AGCGGTATCA GCTCACTCAA AGGCGGTAAT 1920

ACGGTTATCC ACAGAATCAG GGGATAACGC AGGAAAGAAC ATGTGAGCAA AAGGCCAGCA 1980

AAAGGCCAGG AACCGTAAAA AGGCCGCGTT GCTGGCGTTT TTCGATAGGC TCCGCCCCCC 2040

TGACGAGCAT CACAAAAATC GACGCTCAAG TCAGAGGTGG CGAAACCCGA CAGGACTATA 2100

AAGATACCAG GCGTTTCCCC CTGGAAGCTC CCTCGTGCGC TCTCCTGTTC CGACCCTGCC 2160

GCTTACCGGA TACCTGTCCG CCTTTCTCCC TTCGGGAAGC GTGGCGCTTT CTCATAGCTC 2220

ACGCTGTAGG TATCTCAGTT CGGTGTAGGT CGTTCGCTCC AAGCTGGGCT GTGTGCACGA 2280

ACCCCCCGTT CAGCCCGACC GCTGCGCCTT ATCCGGTAAC TATCGTCTTG AGTCCAACCC 2340

GGTAAGACAC GACTTATCGC CACTGGCAGC AGCCACTGGT AACAGGATTA GCAGAGCGAG 2400

GTATGTAGGC GGTGCTACAG AGTTCTTGAA GTGGTGGCCT AACTACGGCT ACACTAGAAG 2460

GACAGTATTT GGTATCTGCG CTCTGCTGAA GCCAGTTACC TTCGGAAAAA GAGTTGGTAG 2520

CTCTTGATCC GGCAAACAAA CCACCGCTGG TAGCGGTGGT TTTTTTGTTT GCAAGCAGCA 2580

GATTACGCGC AGAAAAAAAG GATCTCAAGA AGATCCTTTG ATCTTTTCTA CGGGGTCTGA 2640

CGCTCAGTGG AACGAAAACT CACGTTAAGG GATTTTGGTC ATGAggggtg aagcccttca 2700

gctaacccaa atcAGAAAGC GGGTGGGGAA AAACAACTGG CCGCCACCGG GGAGAACCCG 2760

TGGCCACCAG CGGGGAGAAC TACTGGCCGC CTACGGGGAG AACTTCGTGG CCATTGACAa 2820

gtgcggcgta tgacagatga tgacgtgtgg GGTACCgtgt ttgtgcaggt ttcgcgtgtt 2880

gcagtccctc gcaccggcac ccgcagcgag gggctcacgg gtgccggtgg gtcgACTAGT 2940

TCAGTGATGG TGATGGTGAT GCTCGAGAGA TCTttagctg aagggcttca ccccttgcc 3000

ggtggtggct tcggtgaggc ggaaggagtc gccgtcggtg acgacgacgt gcgcatggtg 3060

caggagccgg tcaacggtgg cggtggccaa tgttttgggc atgatctcgt cgaagccgga 3120

ggggtgcagg ttgctcgaga cggcaatcgc tcggcgttcg taggcagcgt cgacgaggcg 3180

atagaatcct tcggcggcgt cttcagagac cgggagcaga ccgatgtcgt cgacgatgat 3240

caggtcggag cggatgatcc tcgtcagggc tcgggcgatg gagtcgtcgg cacggtgccg 3300

gcggaccagg gcaccgaggt cttcgatggt gaaccaagac accgccagcc cagcttcgac 3360

tgcggcctgc ccgagtgcct cggtgaagtg agacttgcca gttcccgaag gcccacagac 3420

gcagaaacat tcacgccgcc cgacccactc cagggtgcgg agcgcgtcct gggtggcgcg 3480

gggaatcgac gatttcgttt cgtcccagtc gccgaaggtc ttgccggtag ggaaacccgc 3540

gcgtttgcgg cgagtgtgca ggttcgcccg gtcccggccg gcggcttcct cagcgaggag 3600

aacccgcacg acttcggcgg ggtcccagcg ttgggctttc gcggtcggga tgatatcgct 3660

caaggacctg cggatatggg gcagtttcag gcgtttagtc aggtcgatcg cctcggccaa 3720

cgggtcgccg ttagcgctgg taacagtgcg caaaggggtg ctcaTCAGGC GTCATCGCTT 3780

TCATCGTCAG TAGTGGCTGG GCTGGCCGGC GCTGGTGTCT GGCCGAAGGT GGACCAGGCG 3840

GCGGTGCCCG GTTGCAGAGT GTGGTTCTCG CTGCGACGGA CCGGATCAGC GTGTTCGTGG 3900

GTGACTTGGT AGTCGAGGAT CGAGATCAGG TCCCGATCGG CGAACCGGCC GGTCATCGCC 3960

GCCGTCCCCA ACGCTCTGTC GACTGCGTCC GGGGAGTGCA GCTTCGAAAG GTCCACGGCC 4020

TCGGCCATCT TCGCCCTGAT GCGGCGCGCT CCCGCGGCGG CAGCCTCGAC CAGCCAGCTG 4080

TTGGCACCCT GGCCCAACTG TAGGAAGGCG AGCTCGGCGG CGGTACGAGG CTTGGGCACC 4140

CGGTCGGCGG CGTCGCGGTC TGGGCGGGGT GGGTAGTGGT CGTTGTCGAG GACCGGTGAA 4200

CCGGGTTGGC CGCGGCGGTG CCGCGCGACT TCGACTGCAC CGCCATCATT GTCGACGGCG 4260

GTCACGATCA GCTCCTCACC ATGGAAGCGG GCCCAGACCC GAGTGTCGAT CAACGTGTGC 4320

GGCACCGAAT ACCGCACCCC TTCCACCGAG ATCGTCGACT CCCAGTTCAC CCGCCGGGTA 4380

GTGCCGAACA CCGCGGTGAA CGGAGATTTC GGCAGCGGAT GCAGCCGCTG CAACTCTTCG 4440

GCCAGCCGCT CGACCGGCCG CCGCCGAGTA GAGCTGTGGA CGCGGGTGTT GACGTCGGTG 4500

CAGAACTGCC GGCAAGCCGA CTCGAGCTCA CCGAAAGAGT GATACTGCTC GCGCAGGTTC 4560

ACCTCCTTCG GCAACAGGTC GGCTTTGGCG ATCCGCACCG TCGCCTCACT TCCTCCCTTC 4620

GACTCCGGAT CAGCAGGTAA ACACGTCCGC AGCGTCGTGC CGTAATGCCG TGCCACCTCC 4680

ACGATCTCCG GATTGCGGAC CGCGATCCCG GCCACGTGAT CGATCGTCGC GGTCTTCTCG 4740

TTGTCGGTGA GGACATAGAC CGGAACCCCG CCCAAGCGGC GGAATGTCGC ATCCAAACAC 4800

GCCGCCACCG TCGGCAATGT CTTGTCCCAG ATCGGGATCA CCACCCGAAA TCGCGACCAC 4860

GCCAACCATG TGCAGAACAA CGTCGTCTTC CGGCCGGAAA TCGTTGGCCC GTCACCGAAG 4920

TCGTACTGTA ACCACAGCCC AGGTTCGGTT ACCCACGGCC GATACACTCG CCGCCGACCA 4980

GCCCGAAACT GGGCCTTCGC CTCCGCGACC GTCCGGCGCG TGGTCCGCTC ACCGCCGGTG 5040

AACCCCAGCG CGACGATCCG TTCGTGGATG ACATCCGCGC GGACCTTGCC CTGCGAGCGG 5100

ACCACCAACT CCTCGATCTT CGGCAGAAAG TCGTCGATCG CCCGCGCTCG ATGCCGCCGG 5160

CGGTCCGGCG GCTGCCCCGC AGCGCGCATC TGCACATACC GGGCCACCGT GTGGTGATCA 5220

CACCCAGCCA GCTCCGCCGC CGCCCGATAA CTACCCGTGA GGTCGTACGC CTCCAAAATT 5280

TCCATGATCT CCCTGCTCGA TTTcatatgT ATATCTCCTT CTTAAAGTTA AACAAAATTA 5340

TTTCTAGACg ccgtccaTTA TAcctcctca cgtgacgtga ggtgcaagcc cggacgttcc 5400

gcgtgccacg ccgtgagccg ccgcgtgccg tcggctccct cagcccgggc ggccgtggga 5460

gcccgcctcg ata 5473

Appendix B: Primers for Mutant Sequencing

Initial Sequencing Primers

pTNR 199 – 222 Reverse: TGAGTGCTTGCGGCAGCGTCTAG

pTNR 2611 – 2635 Forward: GATCCTTTGATCTTTTCTACGGGG

RMP 2.31 Primers (Forward)

RMP 2.31-359 CATTGCCGCGATGGACAC

RMP 2.31-355 AGTGGCAAAAACAGAGTCCG

RMP 77.23 Primers (Reverse)

RMP 77.23-483 GTGGCCGACGAGGTACTTG

RMP 77.23 – 328 ATGAATGCGTTTGCTCGGTG

Appendix C: Non-producing Mutant Raw Sequencing Data

Sdfsfsfds indicates where two pieces of sequencing data merged

RMP 2.31 (Forward)

```
ACGCCCGTTC GAACAGCTCC CATGTTGTCT CGAGCAGTTC GCCTGCCACG TCATCGACCT       60
GCGCGGAATG CGACGCGTAA TCGACGTCGA TCCGCCGCGC CCGTACCTGC ACTGCAGCAC      120
AGTCGGCCAC CAGCTCGTCG AGCGCCGAGA CCTCGCCGGA AATCACCACG CTCTCAGGGC      180
CATTCATTGC CGCGATGGAC ACGCGCGCAC CGAAGCGGAC GATCAGCGCT ATCGACTCGC      240
TCCGATGGCA ACGCGATCGA CACCATGCCG CCATGGCCTG CGAGCCGCTG CGCGATGACC      300
CGACTACGGG CCACCACCAC CTTCGCCGCA TCCGCAAGAC TCAACGCGCC CGCCACGCAC      360
GCCGCTGCAA TCTCACCTTG TGAATGACCA ACCACCACAT CCGGTGTCAC ACCACACGAT      420
TCCCACACCG CAGCCAACGA CACCATCACC GCAAACGACA CCGGCTGAAC AACATCCACC      480
CGATCGAGCA GCTCCTGCGG CCCCGCCACA AAATCCAGCA AAGACCACCC GGTCAACGGA      540
TCCATCACCG CAGCACACTC AGCCAACCGA GTGGCAAAAA CAGAGTCCGA CACCAACAAC      600
TCACGACCCA TACCCACACC ACTGGGCACC CTGTTACCAG GAAACACAAA CGCCACCCTG      660
CCGTTTACCC GGCTGCCGGT CAGCCCGGGC GCACCGGCCG CCACATCGGT GAGCCCACGG      720
ATCAGATCGT CTACGGCACC GGCCAGCACG ACCGCGCGGT GATCCAGCGC CGCGCGCCCC      780
CTTAGCAGCG CCCCGGCGAT CTCCGCGAGC GCGACATCCC CCGTCGCACC CAGGAAATCG      840
ACAAGCCGCG CAGCCTGCTC TGACAATGCC GCAGGCGTCA CGGCCCGATA TGACGAACGG      900
CAGCACCGAC CCCGACGGGC CAGCAGTTTC GACCGGACGG TCGTCGTGCG GCGGCTGTTC      960
GATGATGACG TGGGCGTTGG TACCAGTCAC CGAGAACGAC GACACCCCGG CCCGGCGCAC     1020
CCGGCCCGTT ATCCGGCCAC GGCTCAGCCT CGACACGCCA GCCGCGCGAC GCCCGCCGAC     1080
CAGTCGATAT GTGGGCTCGG CGCATCGATG TGCAGCGTCA TTAGGCACCA CGCCATGGCG     1140
CATCGCCTCG ATCATCTGGA TGACCCCACC CATGCCTGCG GCCGCCCTGT GTGTGACCCG     1200
ATACTNCCAT TTNGAATCAA TTCCNAAGTA ACCAACTGGG GCCCCCCG                  1248
```

RMP 2.31 (Reverse)

```
CACCCAGCGT TGCCGGGTAC CGAATCTGAC TCACAGCAGG TGGGGAACTA TCTGGCCGTA       60

CGCGGGGAGA AACCATGGCC ACAAAATGCG CCATAAATGG CCGCCAGTGG GGAGCTTACG      120

ATGGCCGTCG ACACGGGCGG GTACCGACCC GTTGTCGTTG CGGGGCAGCC GGACCGGTGT      180

CTTTGCCGGG GTGATGTACC ACGACTACGC GGCCGGGCTC GGCGGTGCCG CCGACGACCT      240

CATCGGCTAT GTCGGCATGG GTAGTTCCGG CGGCGTGGCC TCCGGCCGCA TCTCCTACAC      300

CTTTGGGCTC GAAGGTCCCG CGGTCACGAT CGATACAGCC TGCTCGTTTA TGCACTGGTG      360

GCGCTGCATC TCGCAGCACA GGCGGTGCGC GCCGGGGACT GTTCGCTGGC GGTGGCCGGT      420

GGTGTCCCGG TCATGTCNAC CCCGGGATCG TTCATCGAAT TCAGCAAGCA CCGTGGCCCG      480

GCACCGGATT GGGCNCNGCA AATNCTTCGC CGTGCCCCAA TTGGTTGCGA ACAGGGTNGA      540

AANGCCTTNG GGG                                                         553
```

RMP 77.23 (Forward)

```
CGTTGCCGGG ACCGAATCTG ACTCACAGCA GGTGGGGAAC TATCTGGCCG TACGCGGGGA       60

GAAACCATGG CCACAAAATG CGCCATAAAT GGCCGCCAGT GGGGAGCTTA CGATGGCCGT      120

CGACAGCGGC GGGCGACGAC GACCCGGTCG ACCTCGTCGA CGCTGCGATC TGGGGCCTCG      180

TTCGTGCGGC GCGGGCCGAG TTTCCGGACC GATTGCTGTT GGTCGACCTC GACGCGCAGC      240

CCAGCTCAGC CTTGATCGCG GCGCTCGCAC AGACGGACGA GCCGGAGCTA TCCGTGCGCG      300

GTGGCGAAAT TCGGGTGCCG CGGTTGATGC GGGCCGAGTT GCCCGCGGAG GCCCGTACCT      360

CCTGGAACCG GCAGGGCAGC GTGCTGATCA CCGGCGCGTC CGGCGCACTG GCCACTGTGA      420

TCGCCCGGCA TGTGGTGGTG GAACACGGTG TCCGTCACCT GATTCTAGTG AGTCGTGGTG      480

GCGTGGCCGA CGAGGTACTT GCCGAGCTGG CCGATCTCGG TGCCGGAGGT GACCTCGGCG      540

CGATGCGATG TCACCGATCC GGACCGCGCT TGCTGCGCTC ATCAACGGGG GTGCCTGCCC      600

NACACCCCAC TGTANCGGTC NTGCACCCCC CGGGGTGTNC CNACAACGTG TCTTCCATNA      660

CCCCACGTGA TC                                                          672
```

RMP 77.23 (Reverse)

CGTTGCCGGG ACCGAATCTG ACTCACAGCA GGTGGGGAAC TATCTGGCCG TACGCGGGGA 60

GAAACCATGG CCACAAAATG CGCCATAAAT GGCCGCCAGT GGGGAGCTTA CGATGGCCGT 120

CGACAGCGGC GGGCGACGAC GACCCGGTCG ACCTCGTCGA CGCTGCGATC TGGGGCCTCG 180

TTCGTGCGGC GCGGGCCGAG TTTCCGGACC GATTGCTGTT GGTCGACCTC GACGCGCAGC 240

CCAGCTCAGC CTTGATCGCG GCGCTCGCAC AGACGGACGA GCCGGAGCTA TCCGTGCGCG 300

GTGGCGAAAT TCGGGTGCCG CGGTTGATGC GGGCCGAGTT GCCCGCGGAG GCCCGTACCT 360

CCTGGAACCG GCAGGGCAGC GTGCTGATCA CCGGCGCGTC CGGCGCACTG GCCACTGTGA 420

TCGCCCGGCA TGTGGTGGTG GAACACGGTG TCCGTCACCT GATTCTAGTG AGTCGTGGTG 480

GCGTGGCCGA CGAGGTACTT GCCGAGCTGG CCGATCTCGG TGCCGGAGGT GACCTCGGCG 540

CGATGCGATG TCACCGATCG GACCGCGCTT GCTGCGCTCA TCAACGGGGT GCCTGCGCGA 600

CACCCACTTG TAGCGGTCGT GCACACCGCG GGTGTGCTCG ACGACGGTGT GCTGTCCGCA 660

TTGACCCCGC AGCGGTTGGA TGCCGTCTTG CGCCCGAAGG TCGATGCCAC GTTCCACCTC 720

CACGAACTCA CCGCCGCGCT CGACCTCGAA GCTTTCGTGC TGTTCTCCTC GGCTGCAGGT 780

ACTTTCGGCG GTGCGGGACA GGGCAACTAC GCCGCCGCCA ATGCATTCAT GAATGCGTTT 840

GCTCGGTGGC GGCGCGCGCA GGGAATGCCC GCGTCGGCGC TGGGGTGGGG CCAGTGGGCG 900

GGCGGCGGCA TGGCGGCGGG TCTCGATGAA ACGAGCGCGC GGCGGCTGCG GCGGGCCGGT 960

GTTCTGCCGA TCTCGGCATC GCTCGGCACG ACCCTGTTCG ACCTAACTAC GAATGCCGAC 1020

GACGCGGTTC TGCTGCCAAT GCGACTCGAT CTCGCGGCGA TGCGGTCGGA AACTCCGCCG 1080

GCGCTCCTGC GCACCCTCAT CGACGCCGAG GCATCCTCCG CTGCAATCGG TTTCGCTTCG 1140

GAGCTCTCGG GTGCGACCGC GGCGGAGCAG CGGCGACTGG TCGTGGACCT GGTCCGCACC 1200

CAAGTCTCGG CAGTGCTCGG ACACCGCGGC GCTGCCGCGA TTGATCCGGG CCGCCCATTC 1260

GCCGAGCTCG GATTCGATTC GCTCACCGCG GTCGAGCTAC GGAACCGGCT CAGCTCGGTG 1320

ACAGAGTTGC GGCTGCCGTC CCACGCTTGT GTTCGACCAT CCCGACGCCG ACTGTGCTCC 1380

CCGAGTCCCT GCTGGACACG GTGCTCNATC GGGCCGTAAC CACGGCATCG TACCCCTCGC 1440

CGCGGGGCCA TCGATAACCT ATCGCCATCT GGCATATTGT NCTCCCCGGG CGTCGGCCNG 1500

AAACTTGGAC NGTGGNGCC 1519

RMP 46.37 (Forward)

```
CGCACGGCCA TCGTAAGCGC CCCACTGGCG GCCATTTATG GCGTCATTTT TGTGGCCATG       60
GTTTCTCCCC GCGTACGGCC AGATAGTTCC CCACCTGCTG TGAGTCAGAT TCGGTCCCGG       120
CAACGCTGGG TGGATGGATC CTCTANACGC GCCGCAAGCA CTCAANGGCT TCTTTTTTGT       180
TGCTGCTAAA GGAACGGAAC ACTAAAAACC AGCCCGCAAA ACGTGCTACC CCGATAATTT       200
```

RMP 46.37 (Reverse)

```
GGTGGGGCGG GACGGAGTTA CCGAGGGCAG CGGTGGCCAG GGACCGAATC TGACTCACAG       60
CAGCGTGGGG AACTATCTGG CCGTACGCGG GGAGAAACCA TGGCCACAAA ATGCGCCATA       120
AATGGCCGCC AGTGGGGAGC TACGATGGCC GCGACCGGTC AAGGGCCACG AAGTCCCCGA       180
GCGGCAGAGG CCCCCCGGGG CCCGGTCCCC GGGGCGCGTG TTCCCCCCAC ATGGTACGAA       240
GGCCACCCTA CCAACCTACT CATACNCACG ACTAACCNAA NANGACTCTT TTTTTTTTAT       300
NNNCTATCA                                                               309
```

RMP 46.43 (Forward)

```
TCGACGGCCA TCGTAAGCTC CCCACTGGCG GCCATTTATG GCGCATTTTG TGGCCATGGT       60
TTCTCCCCGC GTACGGCCAG ATAGTTCCCC ACCTGCTGTG AGTCAGATTC GGTCCCGGCA       120
ACGCTGGGTG GATGGATCCT CTAGACGCTG CCGCAAGCAC TCAGTTGGTC TTCAAGGGCT       180
GCTAAAGGAA GCGGAACACG TANAAAGCCA GTCCGCATAA ACGGTGCTGA CCCCGGATGA       240
ATGTCAGCTA CTGGGCTATC TGGACAAGGG AAAACNCAAG CGCAAAAAAA AAGCAGGTAA       300
CTGCAGGGGC TACATGGGCN ATACTANACT GGGCGGTNAT TGACTCAACA NGAATTGNAN       360
GGGCTTTCTT C                                                            371
```

RMP 46.43 (Reverse)

```
CTAGGATCAC TCCACCCAGC GTGGCCGGGA CCGAATCTGA CTCACAGCAG GTGGGGAACT       60

ATCTGGCCGT ACGCGGGGAG AAACCATGGC CACAAAATGC GCCATAAATG GCCGCCAGTG       120

GGGAGCTTAC GATGGCCGTC GACACTTGTC AATGGCCACG AAGTGCTCCC CGTAGGCGGC       180

CAGTAGTTCT CCCCGCTGGT GGCCACGGGT TCTCCCCGGT GGCGGCCAGT TGTNTTCCCC       240

ACCCGCTTCT GATCGGGTTA GCTGAAGGGC TCACCCCTCA TGACCAAAAT CCCCTTAACG       300

TGAGTNCGTC CACGACGTAC ACCCCGTACA AAAAATCCAA GATCTCTGTG TATTTATTTT       360

NANTCCTAAT CGCGCTGCAA ACAAAAACCA CGTTACCACG TGTNGTTGCA GATNACACTC       420

CAAGTACGCN AAC                                                          433
```

RMP 71.3 (Forward)

```
GTGGAACGTA AAACTCACGT TAAGGGTATT TTGGTCATGA GGGGTGAAGC CCTTCAGCTA       60

ACCCAAATCA GAAAGCGGGT GGGGAAAAAC AACTGGCCGC CACCGGGGAG AACCCGTGGC       120

CACCAGCGGG GAGAACTACT GGCCGCCTAC GGGGAGAACT TCGTGGCCAT TGACAATCAG       180

TCTGCCGATG ATACTCTCGT CTGAATCCGG GGTATACAGC GCCGCGGAAC GCCTTTTGCA       240

GTCGCCCCCA ACCCGAAACG CTTTTGGAAA TCCAACTGTC TGTTTGCGTG ANACCCNCGC       300

CCCGAAACGA CAGAGTGGAC AGTACCGCCC GTTCCACCGC ATTGCCATCA GACGCATGTA       360

TTCGTGACGT GTTCGGGAAA AATTTCCCAC ACTTTCGGGT TCACGGGGAC AGCAGAGCCC       420

GTTTGACGAA CACCCCCGTT GCANTAGGGC CCCGAAGCAA GTTGCTCCGA TGACTTCGAA       480

GATCGGATGC CAGATGCAAT GGGATCCAGC CCCCCACTCC GTAGGTAGTT TGCCCGGGGC       540

GGGCGTTGGG CCACACAGNT ACACTCTTNG GCTCCTTTAG CGCGTCTGCT CGGGACTGCC       600

CCATGAAAAC GTTTTCTCGT CCATGGACNA AGACTCCNTC GCCCTTACCG ATCAGNCTAC       660

AGCGAGGCGA TGTGACTCGA TGCCATACTC TGCAAGTTCG CTCCGCCATT AGCATCGNCG       720

TTTGANGAAT ACCCTCCGNC CGGGGTAGTG GAATAAAACC AANCGAAACC GCCCCAACGC       780

CCTGGGCCTG GAATANGCCA AGAGGGGAAN TGNTCAAAGG AACCACCTTG GGCACTTCCA       840

CCTCGAGCCC CGTATTAAGG TTTGAANTTA TTGGAGCTGA AAAAGCCTTC CCGGGGTGGG       900

CCCCGCCAAG CNTAGGGTCC CCCCTAACTT CAGGTCAGGT CCGCCGAAAG CCAAGTCATT       960
```

```
TTCCCGCCTA AAGAAATCGG TTAGTGGAAA AGTTTGGGCA CGCCCTGGAA CAGGTTGGTT       1020

TGTTTCACCC TTCCCCGGGT TAAATTTGGG GAATTCATCT CCAAAGNGTT CAAAAGGTTG       1080

CAACGCAGGG TCCGGGCACA GAATTCCTCC CGGGCGGGTA TTAACTATNC CGTTTCCCGC       1120

ATGTATTACA ANTCCGNTGC GCGAAACAAC CTGCAACCCC CCCCCTTTAC CGGANTCCAC       1180

GTTTTGAATA CCACAAAATA TTCCCACGGG TTTTTGCCCG GCGGAACTTT GTTTGGGGGC       1240

CCCTTCCCCG AAAAAACCCC GAACCAAA                                          1268
```

RMP 71.3 (Reverse)

```
CAGCGTTGCC GGGTACCGAA TCTGACTCAC AGCAGGTGGG GAACTATCTG GCCGTACGCG       60

GGGAGAAACC ATGGCCACAA AATGCGCCAT AAATGGCCGC CAGTGGGGAG CTTACGATGG       120

CCGTCGACAA ATCGTTGCCG GCGTCAGATA TTCGGTCCGT CGTAGTATCC GCGTGCAGCG       180

TCGCGGTCGC TGATCAGCAG CGGGGAACTT TCGGTCGGTC AAGGCTGGAA TCGGTTCAGC       240

TCCTGTCGCC AAAACTCACG GGGAATTGCT CGACCCCATC CCCTCTGGTA CTACATGTTC       300

GGTCATGTAC GCAGCATTAC GGGCTTGCTT GGGGCGCTGC CGTACGCGTA CCATGACGCG       360

AGCATGGCAG CACTCCGCTC TCCCAGCGCA TGCAGTGCAG ACGAATAGCG GTGCGNCTCT       420

ACGCCTGTCG GTTTCCGGGG TGGTGCTGTG CCATACAATT NGCACAGATC CNTCTCGTAC       480

TTTAGTGCGG NCAGCATAGA CAANGCANAT TANGCAATGA NTNGCCTACC CCGGTTCGAT       540

TTATTGGNTG AAGGGGTTNT CGGNTGNTAA ACCANCGNCN GCNCGGANGG AAAATGGCCC       600

CACNTGNGAN GCNGTGGGCC NANGTGGGNG CNTTGGNCCG GGGGNACCTT CCNCTTGGNA       660

GGTTTATCCC AGGACNTANC NAGNTCTTTT TTTCCNCTNC GGGGGTATGC NGTTCCNGTG       720

GGGGGGAAAA GNGANCCCCC CGCGCGAGGC TTCCGGGTCC GGAAAAACCC GGTTTGGGTC       780

CGANTGTCCC CCCAGAAAAT TTCCACGGAA AAAGCCCCCT TGCCCCAACC CC               832
```

RMP 71.7 (Forward)

Sequencing data was not good.

RMP 71.7 (Reverse)

```
CCGGGGNCAC GGGACAGCGA ATTCGCATCC CACCCAGACG TTGGCCGGGA CCGAATCTGA       60

CTCACAGACA GGTGGGGAAC TATCTGGCCG TACGCGGGGA GAAACCATGG CCACAAAATG       120

CGCCATAAAT GGCCGCCAGT GGGGAGCTTA CGATGGCCGT CGACCTGTCA ATGGCCCGAA       180

GTCTCCCCGT AGCGCAGTAT CTCCCGGGGC CGGTCCCCGG GCGCATGTTC CCCCCGCTTC       240

GATGGGTAGC GAAGGCTCAC CCNCATGACC AAATCCCTTA ACGTGAGTCT CCCGTTTCCA       300

CGTGACTCGT TACAGACCCC CCGCCACAAA GAAACCATTT CCCCAAGCGA AATCCTTTTT       360

TTTTTTAA                                                               370
```

RMP 82.5 (Forward)

```
GGCGCCGGGA CCGTTGCGAC CCCGGGAAGC AAAACCCGTC AAGGGATGTC CCGGTCCTGA       60

GGGGTGAAGC CCTACAGCTA ACCCAAATCA GAAAGCGGGT GGGGAAAAAC AACTGGCCGC       120

CACCGGGGAG AACCCGTGGC CACCAGCGGG GAGAACACTG GCCGCCTACG GGGAGAACTC       180

GTGGCCATGA CAAGTGTCCA CGGCCATCGA AGCCCCCACT GGCGGCCATC ATGGCGCATT       240

TGTGGCCATG GTTCCTCCCC GCGTACGGCC AGATAGTCCC CACCGCTGTG AGTCAGATCG       300

GTCCCGGCAA CGCGGGTGGA TGGATCCCTA NACCGCCCCA AACACNAGGC TCTTTTTGGC       360

TGCAAGAACG AACACTAAAA CATCAAAACG GCNACCNATA TNCACNCGTA TGANGACCCG       420

AAA                                                                    423
```

Rmp 82.5 (Reverse)

Sequencing data was not good.

Appendix D: DNA Sequence Analysis of Cloned Transposon Insertion Sites

RMP 2.31 Blastx results aligning among all bacterial genomes.

Sequences producing significant alignments:

Select: All None Selected:0

Alignments Download GenPept Graphics

Description	Max score	Total score	Query cover	E value	Ident	Acces
beta-ketoacyl synthase [Streptomyces sp. GBA 94-10]	160	574	94%	3e-69	52%	ESP9578
hypothetical protein [Streptomyces sp. GBA 94-10]	160	782	94%	3e-69	52%	WP_0377
beta-ketoacyl synthase [Streptomyces sp. PVA 94-07]	160	571	94%	4e-69	52%	WP_0234
hypothetical protein [Streptomyces sp. PVA 94-07]	159	784	94%	4e-69	52%	WP_0379
NapE protein [Streptomyces roseosporus NRRL 11379]	164	307	89%	1e-67	57%	EWS9029
rifamycin polyketide synthase [Streptomyces roseosporus NRRL 15998]	167	637	92%	1e-67	58%	EFE7306
beta-ketoacyl synthase [Streptomyces filamentosus]	167	636	92%	1e-67	58%	WP_0327
rifamycin polyketide synthase [Streptomyces roseosporus NRRL 11379]	166	776	92%	2e-67	58%	EWS9029
polyketide synthase [Pseudonocardia sp. P1]	152	302	84%	3e-66	56%	WP_0332
HbmAII [Streptomyces hygroscopicus]	165	573	93%	1e-65	51%	AAY2822
hypothetical protein [Amycolatopsis jejuensis]	144	298	87%	3e-65	54%	WP_0342
hypothetical protein [Streptomyces sp. RSD-27]	149	298	92%	3e-65	53%	WP_0428
hypothetical protein [Streptomyces sp. Mg1]	147	655	90%	3e-65	57%	WP_0377
polyketide synthase type I [Streptomyces sp. Mg1]	147	655	90%	3e-65	57%	EDX2617
EbeB [Streptomyces aburaviensis]	145	297	92%	6e-65	55%	AGY6275
acyl transferase [Streptomyces acidiscabies]	148	583	92%	1e-64	54%	WP_0103
hypothetical protein [Streptomyces sp. GBA 94-10]	149	295	87%	3e-64	56%	WP_0377
putative type I polyketide synthase [Streptomyces sp. GBA 94-10]	148	295	87%	3e-64	56%	ESP9578
hypothetical protein [Streptomyces acidiscabies]	153	622	92%	5e-64	53%	WP_040838942.1
NapD [Streptomyces hygroscopicus subsp. duamyceticus]	162	293	89%	1e-63	58%	ABB86422.1
NapD protein [Streptomyces roseosporus NRRL 11379]	166	291	88%	4e-63	58%	EWS90293.1
hypothetical protein [Streptomyces griseus]	139	558	83%	4e-63	57%	WP_037665682.1
polyketide synthase [Streptomyces sp. LZ35]	154	556	91%	7e-63	60%	AFV30247.1
hypothetical protein [Streptomyces sp. PVA 94-07]	143	290	87%	1e-62	55%	WP_037970124.1
putative type I polyketide synthase [Streptomyces sp. PVA 94-07]	143	290	87%	1e-62	55%	WP_023415573.1
acyl transferase [Streptomyces lydicus A02]	137	589	91%	1e-62	55%	AJT68079.1
EbeC [Streptomyces aburaviensis]	148	289	91%	2e-62	56%	AGY62755.1
polyketide synthase [Kitasatospora mediocidica]	149	577	91%	2e-62	54%	WP_035796300.1
polyketide synthase [Kitasatospora mediocidica]	145	288	91%	2e-62	54%	WP_035796303.1
polyketide synthase [Streptomyces scabiei]	137	536	94%	5e-62	51%	WP_013005779.1
hypothetical protein [Streptomyces sp. 769]	133	558	93%	5e-62	52%	WP_039633292.1
beta-ketoacyl synthase [Streptomyces scabiei]	137	287	91%	6e-62	51%	KFG04369.1
hypothetical protein M271_45470 [Streptomyces rapamycinicus NRRL 5491]	151	287	89%	6e-62	56%	AGP60460.1
hypothetical protein [Salinispora arenicola]	154	542	93%	6e-62	56%	WP_020609568.1
hypothetical protein [Salinispora arenicola]	154	541	93%	8e-62	56%	WP_018794789.1
hypothetical protein [Streptomyces sp. NBRC 110027]	133	287	90%	8e-62	50%	WP_045867302.1
Acyl transferase [Streptomyces sp. 769]	127	469	91%	9e-62	52%	AJC56292.1
hypothetical protein [Streptomyces sp. 769]	127	469	93%	1e-61	52%	WP_039640918.1
beta-ketoacyl synthase [Streptomyces sp. NRRL B-24891]	140	1184	96%	3e-61	50%	WP_046501326.1
hypothetical protein [Saccharothrix sp. NRRL B-16314]	139	479	94%	3e-61	50%	WP_033442118.1

RMP 2.31 Blastx results aligning among Rhodocci genomes.

Sequences producing significant alignments:

Select: All None Selected:0

Alignments Download GenPept Graphics

Description	Max score	Total score	Query cover	E value	Ident	Accession
hypothetical protein [Rhodococcus rhodnii]	139	807	92%	9e-62	58%	WP_037266862.1
polyketide synthase [Rhodococcus opacus]	123	430	91%	8e-52	49%	WP_005248622.1
polyketide synthase [Rhodococcus jostii]	125	430	91%	1e-51	49%	WP_011596689.1
polyketide synthase [Rhodococcus opacus]	126	430	91%	1e-51	50%	AII06864.1
type I polyketide synthase [Rhodococcus wratislaviensis]	124	246	89%	1e-51	50%	WP_005566969.1
polyketide synthase [Rhodococcus imtechensis]	124	424	91%	1e-51	49%	WP_007298439.1
polyketide synthase [Rhodococcus sp. JVH1]	124	427	91%	2e-51	49%	WP_009477283.1
type I polyketide synthase [Rhodococcus opacus]	123	425	91%	2e-51	50%	WP_005253390.1
polyketide synthase [Rhodococcus wratislaviensis]	125	429	91%	2e-51	50%	WP_037229223.1
polyketide synthase [Rhodococcus opacus]	125	429	91%	2e-50	48%	WP_012691336.1
polyketide synthase [Rhodococcus sp. BCP1]	125	407	92%	5e-49	49%	WP_029543911.1
polyketide synthase [Rhodococcus sp. EsD8]	125	405	92%	6e-49	49%	WP_044475962.1
Malonyl CoA-acyl carrier protein transacylase [Rhodococcus sp. EsD8]	125	405	92%	6e-49	49%	CCW11690.1
hypothetical protein [Rhodococcus rhodnii]	115	415	83%	6e-49	54%	WP_037266849.1
polyketide synthase [Rhodococcus rhodochrous]	125	405	92%	7e-49	49%	WP_043799606.1
polyketide synthase [Rhodococcus aetherivorans]	124	403	92%	2e-48	49%	AKE90242.1
polyketide synthase [Rhodococcus ruber]	117	410	86%	4e-47	47%	WP_017681581.1
type I polyketide synthase [Rhodococcus sp. P14]	117	231	89%	5e-47	47%	WP_010595362.1
polyketide synthase [Rhodococcus ruber]	115	404	86%	1e-46	47%	WP_040272757.1
hypothetical protein [Rhodococcus rhodnii]	133	448	92%	1e-43	51%	WP_037266855.1
hypothetical protein [Rhodococcus rhodnii]	130	263	84%	3e-43	57%	WP_037266865.1
polyketide synthase [Rhodococcus rhodnii]	97.8	203	89%	5e-39	41%	WP_037266583.1
hypothetical protein [Rhodococcus rhodnii]	94.0	202	87%	9e-39	41%	WP_037266869.1
polyketide synthase [Rhodococcus rhodnii]	100	732	91%	1e-38	41%	WP_037266856.1
polyketide synthase [Rhodococcus sp. UNC363MFTsu5.1]	125	437	92%	9e-38	43%	WP_027499595.1
hypothetical protein [Rhodococcus rhodnii]	93.6	385	91%	9e-24	42%	WP_037266859.1
polyketide synthase [Rhodococcus sp. UNC363MFTsu5.1]	90.1	218	91%	1e-33	41%	WP_037176752.1
polyketide synthase [Rhodococcus sp. P14]	80.9	178	82%	2e-31	38%	WP_010595363.1
hypothetical protein [Rhodococcus rhodnii]	75.5	174	90%	2e-30	40%	WP_037275757.1
polyketide synthase [Rhodococcus erythropolis]	78.6	164	91%	2e-27	38%	WP_037133969.1
polyketide synthase [Rhodococcus fascians]	79.7	157	89%	3e-25	37%	WP_037161130.1
polyketide synthase [Rhodococcus fascians]	79.7	157	89%	3e-25	37%	WP_032364873.1
hypothetical protein [Rhodococcus fascians]	90.9	161	68%	1e-22	43%	WP_032392995.1
hypothetical protein [Rhodococcus fascians]	90.9	161	68%	1e-22	43%	WP_032368751.1
hypothetical protein [Rhodococcus rhodnii]	100	201	69%	2e-21	53%	WP_037258360.1
hypothetical protein [Rhodococcus ruber]	68.9	114	53%	1e-18	40%	WP_040783580.1
hypothetical protein [Rhodococcus rhodnii]	64.7	133	82%	2e-18	37%	WP_037256342.1
hypothetical protein [Rhodococcus rhodnii]	75.5	142	60%	7e-17	39%	WP_037273958.1
beta-ketoacyl synthase [Rhodococcus ruber BKS 20-38]	62.0	107	52%	2e-16	39%	EME58155.1
polyketide synthase [Rhodococcus opacus]	49.7	124	78%	1e-15	32%	WP_005255049.1
mycocerosate synthase [Rhodococcus fascians]	64.7	102	52%	5e-15	40%	WP_037182732.1

RMP 77.23 Blastx results aligning among all bacterial genomes.

Sequences producing significant alignments:

Select: All None Selected:0

Alignments Download GenPept Graphics

Description	Max score	Total score	Query cover	E value	Ident	Accession
Beta-ketoacyl synthase [Streptomyces sp. 769]	219	321	75%	2e-78	60%	AJC56297.1
hypothetical protein [Streptomyces sp. 769]	218	321	75%	2e-78	60%	WP_039640922.1
beta-ketoacyl synthase [Streptomyces violaceusniger]	219	1262	76%	1e-77	59%	WP_014060789.1
beta-ketoacyl synthase [Streptomyces rapamycinicus NRRL 5491]	216	1274	77%	4e-77	59%	AGP57748.1
beta-ketoacyl synthase [Streptomyces iranensis]	217	1251	77%	6e-77	58%	WP_044575058.1
beta-ketoacyl synthase [Streptomyces himastatinicus]	211	593	76%	2e-76	59%	WP_039941904.1
modular polyketide synthase [Streptomyces himastatinicus ATCC 53653]	211	919	77%	3e-76	59%	EFL26030.1
polyketide synthase [Streptomyces rapamycinicus NRRL 5491]	212	781	77%	1e-75	56%	AGP57750.1
polyketide synthase [Streptomyces violaceusniger]	211	791	77%	1e-75	57%	WP_014060787.1
beta-ketoacyl synthase [Streptomyces sp. PRh5]	218	520	76%	2e-75	59%	WP_037965962.1
hypothetical protein [Streptomyces yogyakartensis]	201	576	77%	3e-75	58%	WP_046089936.1
polyketide synthase [Streptomyces sp. PRh5]	211	792	77%	1e-74	56%	EXU62660.1
hypothetical protein [Streptomyces sp. PRh5]	211	792	77%	1e-74	56%	WP_037963359.1
hypothetical protein [Streptomyces sp. AcH 505]	193	1111	77%	1e-64	52%	WP_041991952.1
hypothetical protein [Actinoalloteichus cyanogriseus]	186	298	74%	3e-71	51%	WP_035292392.1
LOW QUALITY PROTEIN: modular polyketide synthase [Streptomyces himastatinicus ATCC 53653]	209	514	75%	4e-71	56%	EFL26024.1
polyketide synthase [Amycolatopsis orientalis]	195	297	76%	5e-71	55%	WP_043836768.1
polyketide synthase [Streptomyces sp. NRRL F-3213]	193	296	76%	9e-71	54%	WP_037810558.1
Beta-ketoacyl synthase [Streptomyces iranensis]	204	702	77%	4e-69	55%	CDR09745.1
polyketide synthase [Streptomyces iranensis]	204	650	77%	4e-69	55%	WP_044581912.1
polyketide synthase [Streptomyces violaceusniger]	187	290	76%	5e-69	53%	WP_014056959.1
hypothetical protein [Actinoalloteichus cyanogriseus]	195	290	75%	6e-69	53%	WP_035276150.1
hypothetical protein M271_31590 [Streptomyces rapamycinicus NRRL 5491]	205	1064	75%	7e-69	55%	AGP57746.1
polyketide synthase [Streptomyces yogyakartensis]	188	288	75%	2e-68	53%	WP_046086293.1
polyketide synthase [Streptomyces sp. PRh5]	204	1096	76%	2e-68	55%	EXU66032.1
hypothetical protein [Streptomyces sp. PRh5]	204	1096	76%	2e-68	55%	WP_037955174.1
polyketide synthase family protein [Frankia sp. CpI1-S]	190	812	77%	2e-68	51%	KJE21273.1
hypothetical protein [Streptomyces sp. NRRL S-1022]	179	288	76%	2e-68	52%	WP_037815833.1
lasalocid modular polyketide synthase [Streptomyces lasaliensis]	190	288	76%	2e-68	51%	CAQ64691.1
putative polyketide synthase [Streptomyces lasaliensis]	189	288	76%	2e-68	51%	WP_032492356.1
modular polyketide synthase [Streptomyces himastatinicus ATCC 53653]	193	288	76%	2e-68	57%	EFL26027.1
polyketide synthase [Actinokineospora enzanensis]	194	286	75%	5e-68	53%	WP_018681307.1
hypothetical protein [Nocardia tenerifensis]	184	354	77%	5e-68	52%	WP_040741098.1
hypothetical protein [Saccharomonospora sp. CNQ490]	189	686	76%	1e-67	52%	WP_037336315.1
hypothetical protein [Streptomyces sp. LaPpAH-202]	191	285	75%	2e-67	52%	WP_037816945.1
beta-ketoacyl synthase [Amycolatopsis vancoresmycina]	181	564	76%	2e-67	54%	WP_033263161.1
hypothetical protein [Streptacidiphilus melanogenes]	182	284	75%	3e-67	53%	WP_042389082.1
polyketide synthase [Streptomyces yogyakartensis]	198	283	74%	6e-67	54%	WP_046086413.1
hypothetical protein [Streptomycetaceae bacterium MP113-05]	162	283	72%	9e-67	53%	WP_023532133.1
polyketide synthase [Streptomyces sp. PRh5]	192	431	77%	1e-66	56%	WP_037955176.1
polyketide synthase [Streptomyces iranensis]	192	424	76%	2e-66	54%	WP_044575056.1

RMP 77.23 Blastx results aligning among Rhodocci genomes.

Sequences producing significant alignments:

Select: All None Selected:0

Alignments Download GenPept Graphics

Description	Max score	Total score	Query cover	E value	Ident	Accession
hypothetical protein [Rhodococcus rhodnii]	176	685	76%	1e-62	54%	WP_037266862.1
polyketide synthase [Rhodococcus opacus]	162	526	79%	2e-62	45%	AII06864.1
polyketide synthase [Rhodococcus wratislaviensis]	162	526	78%	3e-62	45%	WP_037229223.1
polyketide synthase [Rhodococcus imtechensis]	160	491	77%	7e-62	44%	WP_007298439.1
polyketide synthase [Rhodococcus sp. JVH1]	159	495	77%	3e-61	44%	WP_009477283.1
polyketide synthase [Rhodococcus opacus]	160	488	77%	6e-61	44%	WP_005248622.1
type I polyketide synthase [Rhodococcus opacus]	157	485	77%	3e-60	44%	WP_005253390.1
polyketide synthase [Rhodococcus sp. BCP1]	168	479	78%	1e-59	46%	WP_029543911.1
polyketide synthase [Rhodococcus jostii]	152	485	77%	4e-59	45%	WP_011596689.1
polyketide synthase [Rhodococcus opacus]	159	519	78%	5e-59	44%	WP_012691336.1
hypothetical protein [Rhodococcus rhodnii]	136	416	82%	5e-55	48%	WP_037266855.1
polyketide synthase [Rhodococcus sp. UNC363MFTsu5.1]	167	455	78%	3e-53	45%	WP_027499595.1
polyketide synthase [Rhodococcus ruber]	162	445	78%	2e-52	47%	WP_040272757.1
hypothetical protein [Rhodococcus rhodnii]	128	262	78%	1e-51	43%	WP_037275757.1
polyketide synthase [Rhodococcus ruber]	159	438	78%	2e-51	47%	WP_017681581.1
polyketide synthase [Rhodococcus sp. P14]	158	436	78%	3e-51	46%	WP_010595363.1
Malonyl CoA-acyl carrier protein transacylase [Rhodococcus sp. EsD8]	151	413	77%	5e-50	48%	CCW11690.1
polyketide synthase [Rhodococcus rhodochrous]	151	414	77%	7e-50	48%	WP_043799606.1
polyketide synthase [Rhodococcus aetherivorans]	151	414	77%	7e-50	48%	AKE90242.1
hypothetical protein [Rhodococcus rhodnii]	127	209	74%	8e-47	47%	WP_037266865.1
beta-ketoacyl synthase [Rhodococcus ruber BKS 20-38]	122	209	80%	1e-46	41%	EME58155.1
hypothetical protein [Rhodococcus ruber]	121	208	80%	2e-46	41%	WP_040783580.1
hypothetical protein [Rhodococcus rhodnii]	158	207	75%	5e-46	50%	WP_037266869.1
type I modular polyketide synthase [Rhodococcus ruber BKS 20-38]	123	205	73%	2e-45	44%	EME67368.1
hypothetical protein [Rhodococcus ruber]	123	205	73%	2e-45	44%	WP_040780804.1
hypothetical protein [Rhodococcus rhodnii]	167	490	76%	1e-42	48%	WP_037266859.1
beta-ketoacyl synthase [Rhodococcus rhodnii]	144	177	57%	2e-37	43%	WP_037266853.1
hypothetical protein [Rhodococcus rhodnii]	148	512	79%	3e-36	47%	WP_037266849.1
polyketide synthase [Rhodococcus rhodnii]	114	668	87%	1e-32	45%	WP_037266856.1
hypothetical protein [Rhodococcus rhodnii]	108	161	75%	2e-32	43%	WP_037266585.1
short-chain dehydrogenase/reductase SDR [Rhodococcus ruber BKS 20-38]	126	126	52%	2e-30	44%	EME57811.1
polyketide synthase [Rhodococcus rhodnii]	87.0	130	74%	3e-23	41%	WP_037266583.1

RMP 71.3

A

cell division FtsK/SpoIIIE [Tsukamurella paurometabola DSM 20162]

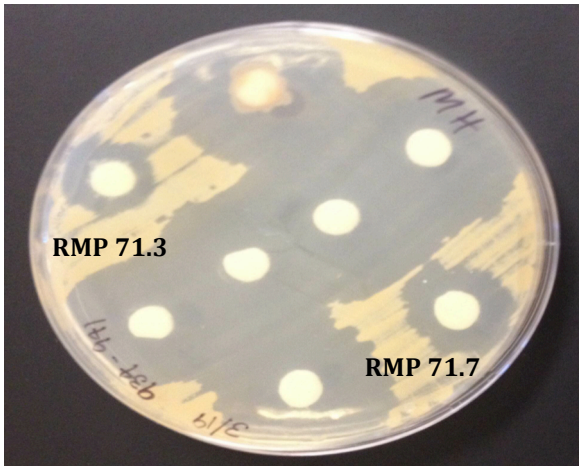
Sequence ID: gb|ADG80837.1| Length: 577 Number of Matches: 1

Range 1: 547 to 577 GenPept Graphics ▼ Next Match ▲ Previous Match

Score	Expect	Method	Identities	Positives	Gaps	Frame
45.4 bits(106)	0.004	Composition-based stats.	23/37(62%)	25/37(67%)	6/37(16%)	-3

```
Query  182  DAARGYYDGPNI*RRQRFVDGHRKLPTGGHLWRILWP  72
            +AA+G  DG        R V GHRKLPTGGHL RILWP
Sbjct  547  EAAKGIVDG------MRGVGGHRKLPTGGHLRRILWP  577
```

B



A1: Amino acid sequence similarity alignment (BLAST). (A) The amino acid sequence of the gene interrupted in the non-producing mutant strain RMP 71.3 was compared with part of the amino acid sequence of the FtsK/SpoIIIE gene from *Tsukamurella paurometabola*. (B) Two of the mutant strains produce the inhibitor compound at a significantly reduced zone of inhibition (RMP 71.3 left, RMP 71.7 right).

RMP 71.7

cell division FtsK/SpoIIIE [Tsukamurella paurometabola DSM 20162]

Sequence ID: gb|ADG80837.1| Length: 577 Number of Matches: 1

Range 1: 551 to 577 GenPept Graphics ▼ Next Match ▲ Previous Match

Score	Expect	Method	Identities	Positives	Gaps	Frame
40.0 bits(92)	0.26	Compositional matrix adjust.	19/27(70%)	19/27(70%)	0/27(0%)	-1

```
Query  188  GETSGH*QVDGHRKLPTGGHLWRILWP  108
            G   G   V GHRKLPTGGHL RILWP
Sbjct  551  GIVDGMRGVGGHRKLPTGGHLRRILWP  577
```

A2: Amino acid sequence similarity alignment (BLAST). The amino acid sequence of the gene interrupted in the non-producing mutant strain RMP 71.7 was compared with part of the amino acid sequence of the FtsK/SpoIIIE gene from *Tsukamurella paurometabola*.

RMP 46.37

A

putative GntR family transcriptional regulator [Rhodococcus erythropolis PR4]

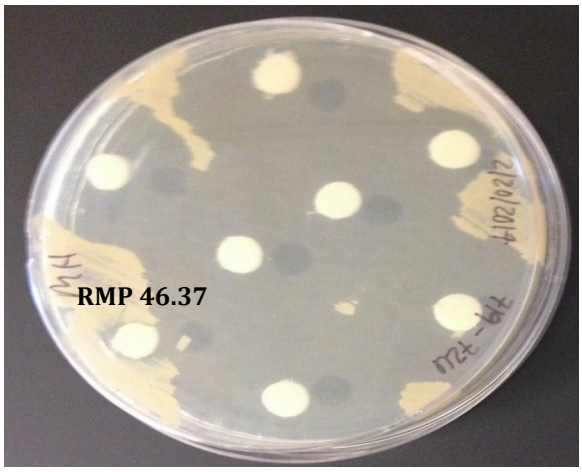
Sequence ID: dbj|BAH31239.1| Length: 489 Number of Matches: 2

Range 1: 478 to 489 GenPept Graphics ▼ Next Match ▲ Previous Match

Score	Expect	Method	Identities	Positives	Gaps	Frame
32.0 bits(71)	0.63	Composition-based stats.	12/12(100%)	12/12(100%)	0/12(0%)	+3

```
Query  48   FCGHGFSPRTAR  83
            FCGHGFSPRTAR
Sbjct  478  FCGHGFSPRTAR  489
```

B



A3: Amino acid sequence similarity alignment (BLAST). (A) The amino acid sequence of the gene interrupted in the non-producing mutant strain RMP 46.37 was compared with part of the amino acid sequence of the GntR transcriptional regulator gene from *Rhodococcus erythropolis*. (B) One of the mutant strains, RMP 46.37, does not produce the inhibitor as observed by no zone of inhibition.

RMP 46.43

A

cell division FtsK/SpoIIIE [Tsukamurella paurometabola DSM 20162]

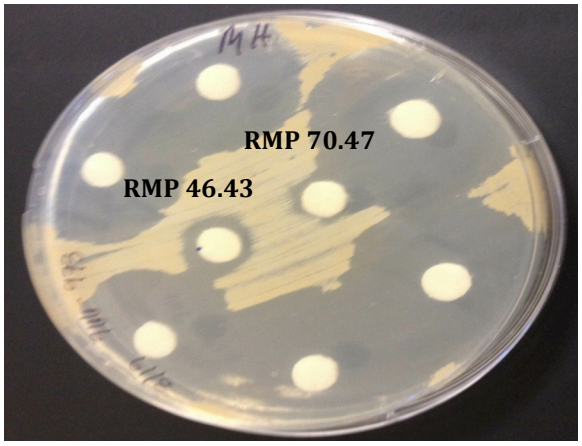
Sequence ID: gb|ADG80837.1| Length: 577 Number of Matches: 1

Range 1: 556 to 577 GenPept Graphics ▼ Next Match ▲ Previous Match

Score	Expect	Method	Identities	Positives	Gaps	Frame
40.0 bits(92)	0.37	Composition-based stats.	17/22(77%)	18/22(81%)	0/22(0%)	-3

```
Query  152  LTSVDGHRKLPTGGHLWRILWP  87
            +  V GHRKLPTGGHL RILWP
Sbjct  556  MRGVGGHRKLPTGGHLRRILWP  577
```

B



A4: Amino acid sequence similarity alignment (BLAST). (A) The amino acid sequence of the gene interrupted in the non-producing mutant strain RMP 46.43 was compared with part of the amino acid sequence of the FtsK/SpoIIIE gene from *Tsukamurella paurometabola*. (B) Two of the non-producing mutants producing a small zone of inhibition against the indicator (RMP 46.43 left, RMP 70.47 right (no sequencing data)).

RMP 82.5

A

beta-glucosidase [Streptomyces fradiae]

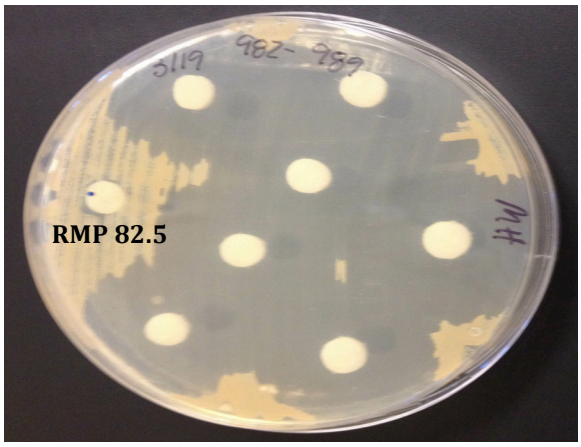
Sequence ID: gb|KDS85082.1| Length: 480 Number of Matches: 1

Range 1: 58 to 97 GenPept Graphics ▼ Next Match ▲ Previous Match

Score	Expect	Method	Identities	Positives	Gaps	Frame
35.8 bits(81)	7.1	Compositional matrix adjust.	17/40(43%)	23/40(57%)	4/40(10%)	-3

```
Query  253  NHGHK----CAMMAASGGFDGRGHLSWPRVLPVGGQCSPR  146
            +H H+        +A +G  D R  LSWPRV+P GG  +PR
Sbjct  58   DHYHRYEEDVTHIAEAGAHDYRFSLSWPRVMPEGGAVNPR  97
```

B



A5: Amino acid sequence similarity alignment (BLAST). (A) The amino acid sequence of the gene interrupted in the non-producing mutant strain RMP 82.5 was compared with part of the amino acid sequence of the beta-glucosidase gene from *Streptomyces fradiae*. (B) RMP 82.5 disk assay that shows that the mutant produces no zone of inhibition against the indicator.

VITA

AMBER LYNNE WARD

Personal Data:	Date of Birth: September 15, 1989 Place of Birth: Jackson, Tennessee Marital Status: Single
Education:	M.S. Biology – concentration in Microbiology, East Tennessee State Univ, Johnson City, TN, 2015
	B.A. Biochemistry, Univ of Mississippi, Oxford, MS, 2013 B.A. Biology, Univ of Mississippi, Oxford, MS, 2013 B.A. English, Univ of Mississippi, Oxford, MS, 2013
	Camden Central High School, Camden, TN, 2008
Professional Experience:	Graduate Assistant, East Tennessee State Univ, College of Public Health, 2014 – 2015
	Undergraduate Research Assistant, Univ of Mississippi, National Center for Natural Products Research, 2010 – 2013
	Teaching Assistant, Univ of Mississippi, Biology Department, 2012 – 2013
Awards:	2015, First Place, Biomedical and Health Sciences, ETSU Appalachian Student Research Forum, Johnson City, Tennessee